TRANSITION STATE SEARCH AND GEOMETRY OPTIMIZATION IN CHEMICAL REACTIONS

By

CRISTIAN E. CARDENAS-LAILHACAR

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1998

To my beloved girls, my wife Alejandra and our daughters Francisca Javiera and Catalina Sofia

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Michael C. Zerner for his support, criticism and teaching. I have been always amazed by his never-ending creativity and enthusiasm: *you can do it Cristian!*; it was always a must. With his very busy schedule, no wonder we couldn't interact more, but somehow he managed to find the time to discuss my progress. He allowed me the time to go ahead with my own ideas, sometimes just to prove to me that I was going in the wrong direction.

In the past few years I have had the opportunity to interact with many people in the Quantum Theory Project, specially in the Zerner research group. Among them I would like to thank Dr. Krassimir Stavrev, Dr. Toomas Tamm, Dr. Marshall Cory, Dr. Guillermina Estiu, Dr. Igor Zilberberg and Dr. Wagner B. De Almeida. Out of the Zerner group, my gratitude for many hours of great science and friendship is given to Dr. Agustin Diz, Dr. Keith Runge, Dr. Ajith Perera, Dr. Steven Gwaltney, and many others that I am probably missing.

I would also like to thank Sandy Weakland, Leann Golemo and Judy Parker from the QTP staff for whom I will hold warmest remembrances.

Friends have always been important for me and my family: Marta and Pradeep Raval, Judy and Marshall Odham, Deborah and Ricardo Cavallino, Sue and Dale Kirmsee, Marcela and Augie Diz, Guillermina Estiu and Luis Bruno-Blanch are friends we will never forget.

My family has always been important to me: my sister Marie-Helene and my

brother Bernard have been good siblings and friends. I always felt lucky to have the parents God gave me, Eduardo and Helene, who gave me so much and ask nothing in return. Few are the occasions that as a friend and a husband I have to express, in a public way, the deep and eternal gratitude that I have for the patience, encouragement, support and love of my wife Alejandra, in whose eyes I saw my future and to whom I gave so little but owe so much. I dedicate this work, my love and my life to her and to our daughters, as they are everything to me.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

ABSTRACT viii

CHAPTERS

1. INTRODUCTION 1

2. REVIEW OF METHODS FOR GEOMETRY OPTIMIZATION AND TRANSITION STATE SEARCH 14

Introduction 14

Geometry Optimization Methods 16

Newton and Quasi Newton Methods 16

The Line Search Technique 18

The Simplex (Amoeba) Technique 20

Restricted Step Method 22

Rational Functions (RFO) 23

Reaction Path Following Method 24

The Hellmann-Feynman Theorem 24

Transition State Search Methods 25

Introduction 25

Simple Monte Carlo and Simulated Annealing Algorithms 26

Synchronous-Transit Methods (LST & QST) 29

Minimax / Minimi Method 32

The Chain and Saddle Methods 33

Cerjan-Miller 35

Schlegel's Algorithm 38

The Normalization Technique or E Minimization 39

Augmented Hessian 40

Norm of the Gradient Square Method (NGSM) 42

Gradient Extremal 43

Gradient Extremal Paths (GEP) 48

Constrained Internal Coordinates 49

The Image Potential Intrinsic Reaction Coordinate (IPIRC) 49

The Constrained Optimization Technique 50

Gradient-Only Algorithms 50

3. HISTORY OF GEOMETRY OPTIMIZATION AND TRANSITION STATE SEARCH IN SEMI-EMPIRICAL AND AB-INITIO PROGRAM PACKAGES 54
Brief Historical Overview 54
NDDO and MNDO 55
MOPAC 56
AM1 56
PM3 56
ZINDO 57
AMPAC (Version 2.1) 58
GAUSSIAN 94 58
HONDO 60
ACES II (Version 1.0) 61

4. THE LINE-THEN-PLANE MODEL 62
Introduction 62
The Line-Then-Plane (LTP) Search Technique 64
The Algorithm 66
Minimizing in Perpendicular Directions: Search for Minima 68
Projector Properties 70
LTP Convergency 71
The Step 73
Default Step 73
Updated Step 74
Newton-Raphson-Like Step 74
Hammond's-Postulate-Adapted LTP Methods 75
Introduction 75
Hammond-Adapted LTP Procedure (HALTP) 76
Restricted Hammond Adapted LTP (RHALTP) 77

5. GEOMETRY OPTIMIZATION 81
Introduction 81
ARROBA: A Line-Then-Plane Geometry Optimization Technique 82
The Algorithm 85
Minima in Perpendicular Directions 86
Convergency 86
A Proposed Global Minima Search Algorithm 87

6. APPLICATIONS 89
Introduction 89
Model Potential Functions for Transition State 90
The Halgren-Lipscomb Potential Function 90
The Cerjan-Miller Potential Function 90
The Hoffman-Nord-Ruedenberg Potential Function 91
The Culot—Dive—Nguyen—Ghuysen Potential Function 94
A Midpoint Transition State Potential Function 94
A Potential Function with a Minimum 96
Summary of Results 96
The Step 103
Step Size Dependence 105
Hammond and Restricted Hammond Adapted LTP Models 105
Summary of Results 110
Molecular Cases for Transition State 112
Introduction 112
Inversion of Water 112
Symmetric Inversion of Ammonia (NH_3) 116
Asymmetric Inversion of Ammonia (NH_3) 122
Rotated Symmetric Inversion of Ammonia (NH_3) 124
Hydrogen Cyanide: HCN → CNH 125
Formic Acid 132
Methyl Imine 136
Thermal Retro [2+2] Cycloaddition Reaction of Oxetane 141
Hammond Adapted LTP Results 147
Summary of Results 147
Model Potential Function for ARROBA 150
Introduction 150
Model Potential Function 150
Step Size Dependence 151
Summary of Results 153
Molecular Case for ARROBA: Water 153
7. CONCLUSIONS AND FUTURE WORK 158
BIBLIOGRAPHY 162
BIOGRAPHICAL SKETCH 169

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy.

TRANSITION STATE SEARCH AND GEOMETRY OPTIMIZATION IN CHEMICAL REACTIONS

By

Cristian Cardenas-Lailhacar

August, 1998

Chairman: Michael C. Zerner
Major Department: Chemistry

The research presented in this thesis involves the development of procedures for finding transition states in chemical reactions as well as techniques to optimize the geometries that are involved in their calculations. A procedure for finding transition states (TS) that does not require the evaluation of second derivatives (Hessian) during the search is proposed. The procedure is based on connecting a series of points that represent products $\mathbf{P}_i$ and reactants $\mathbf{R}_i$. From these points, conservative steps along the difference vector from $\mathbf{P}_i$ toward $\mathbf{R}_i$, and from $\mathbf{R}_i$ toward $\mathbf{P}_i$, are taken, until the two points coalesce. Although the initial points of the set, $\mathbf{P}_o$ and $\mathbf{R}_o$, represent specifically the product and the reactant, other $\mathbf{P}_i$ and $\mathbf{R}_i$ are determined by minimization in hyperplanes that are perpendicular to $\mathbf{P}_{i-1}$ and $\mathbf{R}_{i-1}$, simultaneously. In order to test the accuracy of the methodology proposed here, the technique has been applied to seven well-known potential functions, and the results compared with those obtained from other well-known procedures. Most methods that search for transition states require an accurate evaluation of the Hessian as they proceed uphill from either product to reactant, or from reactant to product. These procedures are both costly in computer time and in memory storage.

The line-then-plane (**LTP**) methods described here do not need the accurate calculation of the Hessian except for the last step in which its signature usually has to be checked. This particular one point could also be probed numerically. This feature potentially allows the study of much larger chemical systems.

When this LTP technique is applied to molecular reactions, the results compare closely with those derived from the application of other models. The proposed LTP geometry optimization procedure, after being tested in a model potential function, has been used together with the LTP technique to define a general procedure to find the global minimum.

It is shown that, because of the Newton-Raphson nature of the step taken, the ***LTP*** procedures will converge unequivocally to the TS on any continuous surface. The same applies for the minima searching in the geometry optimization procedures.

CHAPTER 1
INTRODUCTION

Since the beginning of human reasoning symbols have played a key role in a colossal attempt to try to describe our universe, the cosmos. The history of chemistry takes us back to the symbols of fire, water, air and earth that were extensively used by alchemists around the 13th century and attributed to Plato's polyhedral symbols. Also involved in this historical perspective are Empedocles of Agrigent (~440 BC), Thales of Miletus (~600 BC), Anaximenes (546 BC) and Heraclitus (~500 BC) who claimed fire to be a basic element. Few of these symbols are still with us. The symbol for fire (heat), $\triangle$, is the only one used in chemistry [1].

But alchemy has evolved into chemistry as the scientific method has replaced the old beliefs such as the transmutation of matter (an idea that still today is among some chemists' inquiries). We are still wondering about the inner secrets of matter. Chemists are confronted with many questions, one of the most important being the description of how atoms are held together in molecules, and how they interact with each other to produce new compounds, that is, how chemical reactions proceed. Quantum chemistry has become a powerful tool to assess such a goal.

As quantum chemists we are interested, in general, in accounting for the properties of excited as well as ground states. In the case of chemical reactions, the aim is to understand and describe the laws of nature that control them. To this end, algorithms are constructed, to reproduce features of a chemical reaction on the computer, and

tested in their *goodness* by calculating observable quantities that are finally compared with the experiment.

The potential energy surface (PES) is the cornerstone of all theoretical studies of reaction mechanisms in relation to the chemical reactivity. Topographic features of the PESs are strongly associated with experimental observations of the chemical reaction. A lowest energy path connecting reactants (R) and products (P) (selected ones) on the surface is a concept that can be associated with the mechanism through which the reaction, theoretically, occurs. The association of these pathways with valleys among mountains is as unavoidable as practical and allows us to understand the feasibility of a reaction.

The maxima along the path related to the reaction mechanism are essential for understanding of the energetics of the processes under study. These particular points, which have been called transition states, tell us about the type of reaction with which we are confronted. An insurmountable mountain along another pathway tells us that the associated reaction connected is not feasible. On the other hand, the presence of two transition states is the theoretical equivalent to competing reactions in a test tube, whereas a shallow minimum may confirm the existence of a postulated intermediate. The variety of reaction mechanisms is enormous, and, consequently it is essential to have a good understanding of the properties that are general and common to all potential energy surfaces.

Chemical reactivity is the main subject of chemistry. The goal is to predict the products that are most likely to be obtained according to the interactions among the participating species. In 1889 Svante Arrhenius initiated the study of transition states by expressing the sensitivity of the reaction rate to temperature through his famous relation [2]. Later, in 1931, with the development of molecular reactivity theories and

particularly with the work of Michael Polanyi and Henry Eyring [3], the goal was to formulate relations for the kinetics of reactions. These theories introduced concepts such as *activation barrier* and *transition state*. With the coming of quantum chemistry, rational techniques for the prediction of molecular structures (geometry optimization) and mechanisms of reactions (transition states) became available. Results obtained today reveal the spectacular degree of refinement that quantum chemical theory has achieved, even on occasions competing with experimental measurements for accuracy.

If we plot all the positions and energies of reactants as they evolve to products, we will obtain a potential energy surface (PES). The TS may be like a volcano between two valleys (for a single transition state) or a rugged mountain range (for more than one TS). This multidimensional surface contains many paths with different mountain passages (energy barriers) through which reactants must move to become products. In the particular path that the reaction follows, the transition state is the point of highest energy between reactants and products. This classical view of transition states has evolved today into a broader definition: the full range of configurations the reactants can take as they evolve to products [4]. This difference is mainly due to how we look at the TS, that is, as one point or a realm of reaction rates in a potential energy surface of 3N–6 dimensions (with N being the number of atoms).

Some reactions can go from reactants to products without passing through a transition state via a minimum energy pathway, making the location of the TS a very unpleasant task particularly for experimentalists. This area of the PES is called a *seam*, that is, a region of the PES that is penetrated by another one. This behaviour happens when the energies of the ground and excited states are so close that the system can bypass the transition state. A representative example of this phenomena is the internal rotation of stilbene [5].

Chemical reactions are classified according to the difference in energy between the Rs and Ps, that is, ***endothermic*** and ***exothermic*** reactions, according to:

$$\Delta E^o = E_P - E_R \tag{1}$$

where $\Delta E^o < 0$ and $\Delta E^o > 0$, respectively. The reactions are studied along a given ***reaction coordinate*** (RC) which tells us how to **step**, the walking direction and the evolution of the reaction from reactants to products (initial and final situation) and *vice-versa*. Somewhere between Rs and Ps, there is a maximum in energy that is unavoidable, the ***transition state***, a very unstable conformation that will transform itself into reactants or products according to the initial conditions of the trajectory [6].

What is the appearance of the transition state? What bonds are broken and formed? What structural changes are occurring in the system and why? Unfortunately, after almost three decades, quantum chemistry still does not have effective algorithms to solve this problem, and even today the most common recipe is to locate and describe transition states from chemical intuition, that is to say, experience.

The usual starting point is to optimize both reactants and products geometries: minima in the PES. The TS is then a maximum situated between them. However, there are exceptions to this scheme [5, 6]. In Figure 1.1 we show the internal rotation of hydrogen persulfide, for which the reactants and the products (trans and cis isomers, respectively) have been fully optimized at a fixed dihedral angle of $\alpha = 0$ and 180°, respectively, as chemical intuition will indicate. In turn, the TS is expected to be located, at a higher energy, around midway between reactants and products. Here the expected TS (located at $\alpha = 93°$) turns out to be a minimum and the initial reactant and product maxima along the reaction coordinate. Which is then a minimum and which a TS?

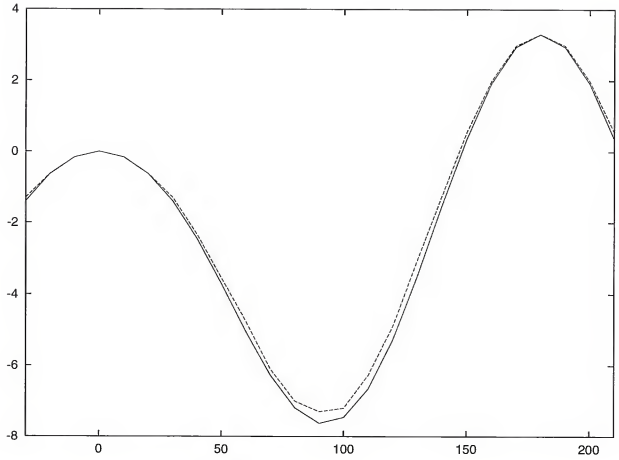


Figure 1.1. Hydrogen persulfide internal rotation. The continuous line represents Ab-Initio (STO-3G) calculations, whereas the dashed line is for the potential energy surface as obtained through a symmetry adapted technique that we developed [5].

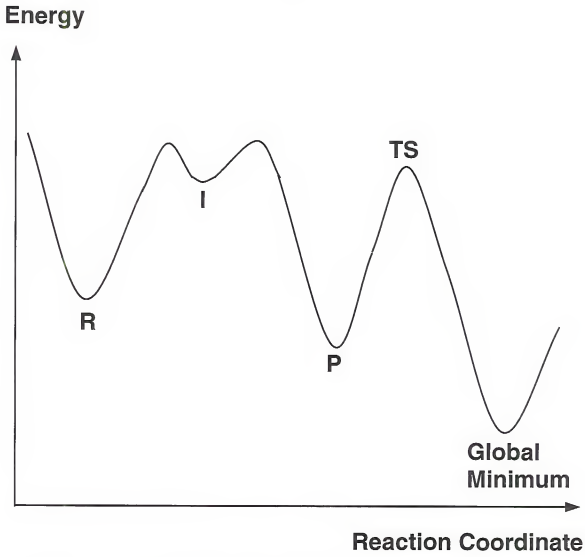


Figure 1.2. The Multiple Minima problem. Energy versus a given reaction coordinate showing local minima, reactants (R), products (P), intermediates of reaction (I), transition state (TS) and the global minimum.

On the other hand, when optimizing geometries the problem as to which is a local and which a global minimum, as shown in Figure 1.2, is still a hazard for big molecular systems. If a given geometry is optimized, chances are that the structure soon

will become trapped in the energy minimum of the potential energy surface closest to the starting conformation. How then is one to find the desired global minimum? One way would be to use brute force, namely to change systematically the value of a givenvariable across the surface. An alternative is to perform a systematic search by covering conformation space with a fine mesh, but this requires too many calculations. Another interesting way to address the problem is to think of minimization algorithms as *cooling* molecular structures to 0° Kelvin, then by a *warming-up* process the system is taken to a higher energy position in the PES, and the search can continue in another region of the N-dimensional conformational energy surface.

This overview simply tells us that we still do not have algorithms that are efficient enough to solve this optimization problem, not to mention the expense in terms of number of iterations necessary to obtain this minimum (when found), that is, computer time. Transition states have only an ephemeral existence (*vide infra*) that lies in the femtosecond scale as shown in the cosmic time scale in Figure 1.3 (if we were able to live 32 million years, the transition state would last only for a few seconds of our lives). Worthy of mention is the time-resolved experimental work of Zewail [7] and collaborators who, by using femtosecond (ultrafast) laser techniques, observed reaction dynamics of small molecular systems. Nevertheless, TS are attainable by quantum mechanics, whereas experimentally they can only be inferred indirectly. This distinction has motivated theoretical chemists to develop new and powerful models to search for transition states [8]. New methods appear frequently in the literature [5, 9–16] and, as we will see in the next chapters, the mathematical tools as well as the models sometimes seem to be directly proportional to the number of scientists devoted to tackling the problem.

But none of these procedures is as yet utterly convincing or generally successful.

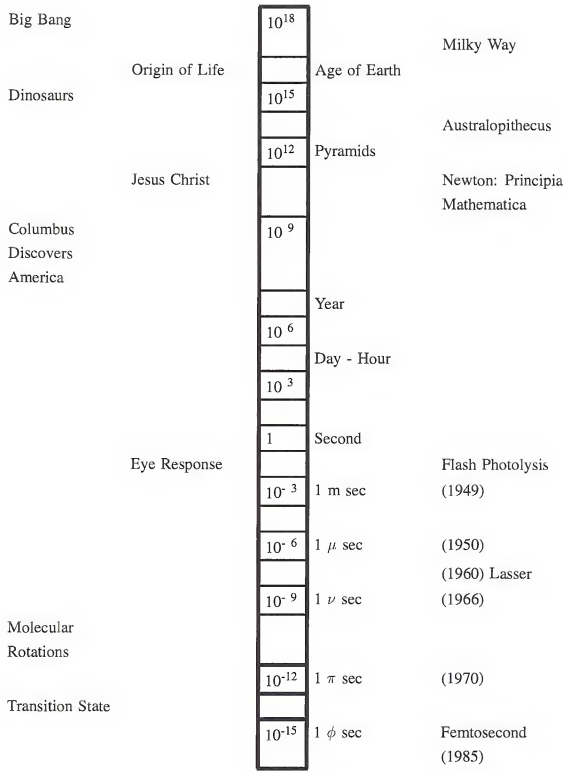


Figure 1.3. Cosmic time scale for transition state (in seconds).

Consequently it becomes very important to examine new algorithms capable of accurately finding minima and transition states. The algorithms usually found in the literature can be divided in two general kinds; those (the cheaper, and usually less accurate, ones) that use only gradients and can give a quick, but rough, idea of the transition state location, and those (more sophisticated ones) that use gradients and Hessians (more expensive but also more accurate, when successful). The most efficient algorithms to find TSs use second derivative matrices which require great computational effort. This fact alone is a powerful incentive to try to develop new procedures that do not require the Hessian.

The determination of TS structures is more difficult than the structure of equilibrium geometries, partly because minima are intrinsically easier to locate and also because often no *apriori* knowledge is available about TS structures.

For a given structure χ_e to be a **TS** of a reaction it must fulfill the following conditions, according to McIver and Komornicki [8]:

- χ_e ***must be a stationary point, which means that all gradients (g) of the energy evaluated at this point must be zero:*** $\mathbf{g}(X_e) = 0$.

- ***The force constant matrix (H) at the transition state must have one and only one negative eigenvalue*** $H(X_e)$.

- ***The transition state must be the highest energy point on a continuous curve connecting reactants and products.***

- ***The point identified as the transition state*** (X_e) ***must be the lowest energy point which satisfies the above three conditions.***

The computation of the energy and its derivatives of the system under study is essential for our purposes. For this, the Born—Oppenheimer Approximation is used which is based on the assumption that, given a molecular system, the nuclei are much heavier than the electrons remaining clamped. As a consequence the kinetic energy of the nuclei is neglected and the repulsion between nuclei is considered to be a constant. This approximation gives rise to the electronic Hamiltonian, which in atomic units for N electrons and M nuclei is

$$\mathcal{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{\alpha=1}^{M}\frac{\mathcal{Z}_\alpha}{r_{i\alpha}} + \sum_{i=1}^{N}\sum_{j>i}^{M}\frac{1}{r_{ij}} \tag{2}$$

where ∇_i^2 is the Laplacian operator (derivatives respect to the coordinates of the i th electron), $\mathcal{Z}_\alpha$ is the atomic number of nucleus α, $r_{i\alpha}$ is the distance between the ith electron and nucleus α, whereas r_{ij} is the distance between electrons i and j . In this Hamiltonian we identify the first term as to be the kinetic energy of the electrons, the second term represents the Coulomb attraction between electrons i and nuclei α and the third term addresses for the repulsion between electrons.

The energy (E) of the system comes from the solution of Schrodinger's equation that we wish to solve using our Hamiltonian operator: $\mathcal{H}\Psi = \mathrm{E}\Psi$, where Ψ is the wave function we use to represent the system under study.

The algorithms that we used for a transition state (TS) search and geometry optimization, as well, are generally based on a truncated Taylor series expansion of the energy

$$\mathrm{E} = \mathrm{E}_\circ + \mathbf{q}^\dagger g + \frac{1}{2}\mathbf{q}^\dagger \mathbf{H}\mathbf{q} + \ldots \tag{3}$$

and of the gradient

$$g = g_\circ + \mathbf{q}\mathbf{H} \tag{4}$$

with $\mathbf{q}$ the coordinates, g the gradient (first derivative of the energy with respect to coordinates $\mathbf{q}$) and $\mathbf{H}$ the Hessian (second derivative matrix of the energy with respect to coordinates $\mathbf{q}$). First derivatives for any wave-function generally can be acquired analytically in about the same time as the energy. Analytical second derivatives, on the other hand, involve at least coupled perturbed Hartree-Fock (CPHF) algorithms. These have, in general, a fifth-order dependence on the size of the basis set, that cannot be avoided if the Hessian is required to find minima, and are imperative to insure the location of a TS.

Today modern procedures try to avoid the evaluation of the Hessian as this is a real bottleneck in the calculation in terms not only of computer time as well as memory storage. Consequently, algorithms that *update* the Hessian (or its inverse), that is, procedures that use a guess of the Hessian and information of the actual and previous structure give a "good enough" estimate of the real Hessian after a few iterations. When the initial Hessian is chosen to be the identity (or other approximate) matrix, the procedure is said to be a quasi-Newton one. On the other hand, "true" Newton procedures are those that use a calculated Hessian.

Update procedures in turn are known to be of two types (see for example [17–19]):

$$\textbf{Rank 1}: \quad G_n = G_{n-1} + W_n$$

$$\textbf{Rank 2}: \quad G_n = G_{n-1} + W_n + V_n \tag{5}$$

where W_n and V_n are corrections to the initial Hessian or its inverse G_{n-1} at cycle n. To rank 1 correspond update procedures such as the one by Murtagh and Sargent (MS) [20], while a popular rank 2 method was constructed by Broyden-Fletcher-Goldfarb and Shanno (BFGS) [21], Davidon-Fletcher-Powell (DFP) [22] and Greenstadt [23].

It is germane to note that rank 2 update procedures can be regarded as being a rank 1 update of an already rank–1–updated Hessian (or its inverse). A great deal of work has been carried out lately in this field: the more recent papers combine rank 2 update procedures [24].

The determination of the minimum energy conformations of reacting species is handled more or less routinely except for very large systems with multiple minima. Transition states are not as easy to find as minima. Moreover, most algorithms that we will describe in the next chapter do not always succeed in the search for transition states because of the following general reasons:

- It is difficult to insure movement along a surface that exactly meets the conditions of a simple saddle point.
- In general, little *a priori* knowledge of the transition state structure is available.
- Wave functions for a TS may be considerably more complex than those describing minima.

Some procedures make a guess of the TS and perform a Newton-Raphson minimization of the energy. Unfortunately this technique is not reliable because it can lead back to R, P or to a TS. Many different algorithms for these tasks are available in the literature, with good reviews found in references [9, 19, 25].

In this work we show methods to find transition states based on a continuous walking of fixed step along a line connecting R and P, assuming that their structures

are known, and utilizing methods to optimize geometries (GOPT) based on the initial and the newly generated structure. No previous knowledge of the TS is necessary.

In chapter 2 we review the existing literature on geometry optimization and transition state search algorithms in terms of advantages and disadvantages, starting with a description of the line-search technique that accounts for parameters used in the majority of the models. We emphasize the disadvantages as they account for costly failures which these procedures suffer.

Chapter 3 starts with a brief historical overview of semi-empirical molecular orbital theories. Next, some semi-empirical and ab-initio program packages are examined in terms of their TS and GOPT capabilities.

Chapter 4 introduces the line-then-plane (LTP) procedure. The algorithm is described discussing its convergence to the TS and how the step should be taken. Alternative algorithms, in the basis of Hammond's postulate, are discussed. We concentrate on some properties of LTP, for example its dependence on the size of the steps in terms of the number of energy evaluations required to find a maxima or minima.

In chapter 5, ARROBA, a new LTP geometry optimization procedure, is presented. The main features of this technique are studied through a model potential function and a molecular example. Finally, an algorithm is proposed to solve the multiple minima problem.

In chapter 6 the LTP technique is tested with some potential energy functions and molecular systems for both transition states and geometry optimization problems.

Finally, in chapter 7 we summarize results, draw conclusions, and set the stage for future systematic work in this area.

CHAPTER 2
REVIEW OF METHODS FOR GEOMETRY OPTIMIZATION AND TRANSITION STATE SEARCH

Introduction

In general, optimization techniques for finding stationary points on PESs can be classified (avoiding details for simplicity), as [17–19]

- Without Gradient,
- With Gradient: Numerical or Analytical, or
- With Numerical or Analytical Gradient and Numerical Hessian.

With the exception of the first method, these algorithms are all based on a truncation of a Taylor series expansion of the energy and of the gradient as was given in Eqns. (3) and (4), respectively, in the previous chapter.

The general scheme is complete when the characteristics of the stationary point are included, that is, zero gradient ($g = 0$). In practice the gradient should be smaller than a preestablished threshold at the critical points which, from Eqn. (4), yields

$$g = -qH . \tag{6}$$

From here the new coordinates are

$$q = -gH^{-1} \tag{7}$$

and the step (**s**) is taken as a fraction of **q**

$$\mathbf{s} = \alpha\mathbf{q} \qquad \forall\alpha \in \Re \;/\; 0 < \alpha \leq 1 \,. \tag{8}$$

The bottle neck for all procedures that search for minima or maxima is the evaluation of the Hessian matrix (**H**), as this is time consuming and requires storage. As mentioned in the previous chapter, the way around this problem is to use techniques that update the second derivatives matrix as MS, BFGS, DFP, and so on. At this point, the techniques used to find the critical points of interest are classified as (exact) Newton-Raphson if no approximations are used for the evaluation of **H**, that is, numerical or analytical second derivatives are used in the search. Procedures that use update techniques for the Hessian are said to be of the Newton-Raphson-like type. Whereas, if the identity matrix (**I**) is used, then we are in the steepest descent regime. This should not be confused with those techniques that use the identity matrix as a starting point to update the Hessian as these, when converged, show a second derivative which not only is not the identity matrix but, moreover, is an Hermitian matrix that is very close to the analytical one.

First derivatives for any wave function generally can be acquired analytically in about the same time as the energy; methods that use numerical gradients typically are not competitive. On the other hand, the analytical evaluation of second derivatives involves, at the very least, a Coupled Perturbed Hartree-Fock (CPHF) procedure and is, in general, an N^5 procedure where N is the size of the basis set. Therefore, second derivative methods are costly even at the single determinant level, and even more so at the CI level. However, a good feeling of the topology of the potential energy surface is needed to locate maxima and minima.

Geometry Optimization Methods

Newton and Quasi Newton Methods

From the Taylor series expansion of the energy (equation (3)) we write for the gradient

$$g(\mathbf{x}+\delta) = g(\mathbf{x}) + \mathbf{H}(\mathbf{x})\delta \tag{9}$$

with δ the infinitesimal step in the search direction. A minimum implies that $g(\mathbf{x}+\delta) = \mathbf{0}$. Thus a relation for the search direction $\mathbf{s}$ can be written as

$$a\mathbf{s} = -\left(\mathrm{H}^{-1}\right)g(\mathbf{x}) = \delta\ . \tag{10}$$

For "a true" Newton method, $\mathbf{H}$ is the exact Hessian matrix, while for the quasi-Newton procedure the Hessian can be a matrix that approximates the second derivatives [26, 27], commonly the identity. The line search parameter a can be determined by a variety of algorithms that reduce the energy along the search direction. This type of procedure will be discussed in detail in the next section.

When the Hessian is positive and exact, the Newton method shows a quadratic convergence to a local minimum if expanded in the quadratic region. The problem is that this procedure requires an explicit calculation of the second derivative matrix, demands the extra computational effort previously discussed and, moreover, is accurate only in the quadratic region.

The quasi-Newton methods depend on information obtained about the Hessian during the search. The gradients calculated at different geometries are used to build an approximate inverse Hessian $\mathbf{G} = \mathrm{H}^{-1}$ using the quasi-Newton conditions (Eqn.

(6)). This is

$$\gamma = g(x + \delta) - g(x)$$

$$\delta = G\gamma \tag{11}$$

as a constraint to obtain a relation to get a matrix product for the search direction **s**. This is

$$G(x + \delta) = G(x) + U \tag{12}$$

and

$$\mathbf{s} = -\mathbf{G}(\mathbf{x} + \delta)g \tag{13}$$

where **U** is a correction to the inverse Hessian **G**. Thus the approximate Hessian is updated every geometry cycle. In this way some time is spent evaluating the approximate Hessian or its inverse, but this task requires a small fraction of the time that the evaluation of the Hartree-Fock gradient takes. One of the most successful relations for updating the Hessian is the one developed by Broyden, Fletcher, Goldfarb and Shanno (BFGS) [21]. Many of the procedures available for finding minima differ on how they evaluate or choose **s**. For computational details on these methods see Kuester and Mize [28].

New procedures to optimize geometry are the subject of study by many research groups and one can see specialized journals frequently publishing works devoted to this problem. Rather than follow a textbook classification, here we summarize only those procedures that have proved to be more stable and are most often used. We will start by summarizing the line search technique, as it is widely used by several minimization techniques. Time-dependent, statistical mechanics and variational procedures will not be addressed nor discussed here (except Monte Carlo techniques).

Reviews of geometry optimization and transition state search methods can be found in the literature [13, 18, 23–25].

The Line Search Technique

The fundamental idea of this strategy is to look for the appropriate displacement along a search direction. Suppose that at a given point X_k we found a search direction determined by

$$s_k = -G_k g_k \tag{14}$$

where k indexes the cycle and the inverse of the Hessian $G_k = H_k^{-1}$ is already updated (for example using BFGS). Then the next point is given by $x_{k+1} = x_k + \alpha_k s_k$. Here the parameter α_k comes from the line search and has a value such that the decrease in energy is reasonable (Figure 2.1). This is

$$E(x_{k+1}) < E(x_k) \qquad or \qquad E(x_{k+1}) - E(x_k) < \varepsilon \tag{15}$$

where ϵ is a given energy threshold. But an exact line search will be able to find the exact value of α_k for which $E(x_{k+1})$ along the line defined by s_k is a minimum. Among the many procedures for performing these calculations, one of the most efficient approaches is to perform an "efficient partial line search," in which a reasonable decrease in the function is obtained when an appropriate value of alpha is selected [29].

The energy along s_k is written as

$$E(x_k + \alpha s_k) = E(x_k) + \alpha g^{\dagger}(x_k) \cdot s_k + \frac{1}{2}\alpha^2 s_k^{\dagger} H s_k \tag{16}$$

and will be lowered provided that the descent condition is satisfied: $g^{\dagger}(x_k) \cdot s_k < 0$. From here the sign for the search direction is selected and the minimum is found by direct differentiation

$$E(x_k + \alpha_m s_k) = \left(\frac{\partial E}{\partial \alpha}\right) = g^{\dagger}(x_k + \alpha s_k) \cdot s_k = 0 \tag{17}$$

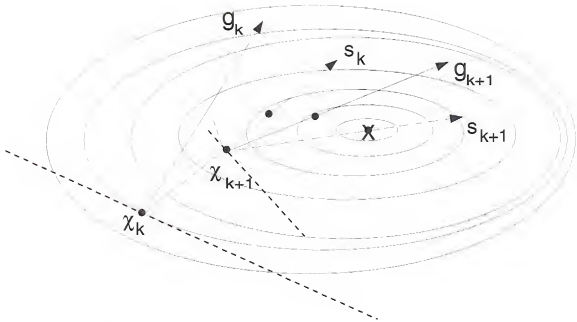


Figure 2.1. Representation of the line search technique in a potential energy surface. g_k and g_{k+1} (continuous arrows) are the directions of greatest slope at points χ_k and χ_{k+1}, respectively, and are orthogonal to the tangents to the surface (dashed lines) at these points. The point on $\mathbf{s_k}$ represents the result of a *partial line search*, whereas on $\mathbf{s_{k+1}}$ the exact line search and the partial line search results coincide.

where the null value characterizes the exact line search. Now the energy at x_{k+1} is required to approach the x_k—$\alpha_k s_k$ value. Then the gradient $g(x_{k+1}) = g(x_k + \alpha_k s_k)$ is evaluated and a left extreme test is performed as

$$\left| \langle g(x_{k+1}) | s_k \rangle \right| \leq -\sigma \langle g(x_k) | s_k \rangle \quad , \qquad 0 \leq \sigma \leq 1 \tag{18}$$

If σ=0 an exact line search is performed, and if σ=1 any reduction of the scalar product $\langle g(x_{k+1}) | s_k \rangle$ is acceptable.

If the left test fails, α_k is too low, then a new value of alpha is estimated as

$$\alpha_k^{new} = \alpha_k + \frac{(\alpha_k - \alpha_k^1)\mathrm{E}'(\alpha_k)}{\mathrm{E}'(\alpha_k^1) - \mathrm{E}'(\alpha_k)} \tag{19}$$

where α^1 is the initial value of the interval (α_k^1, α_k^2) used to examine the kth line search cycle.

Finally the energy $\mathrm{E}(x_k) \rightarrow \mathrm{E}(x_k + \alpha_k^{new} s_k)$ is evaluated and the test repeated with $\alpha_k^1 \rightarrow \alpha_k$ and $\alpha_k \rightarrow \alpha_k^{new}$ until the condition is fulfilled. The partial line search stops and a new search direction is sought.[1]

The Simplex (Amoeba) Technique

Designed by Nelder and Mead [30], this procedure has as one of its main characteristics its geometrical behaviour and that it only requires the evaluation of the energy, having as major drawback the large amount of energy evaluations necessary to find the minimum. The different behaviours, according to possible different topological situations, that the algorithm might have are shown in Figure 2.2.

For molecular systems, a *simplex* is a 3N dimensional geometrical structure with N vertices (in 3 dimensions it is a tetrahedron). In the multidimensional complex topography of a molecular system, simplex requires only 1 (3N-dimensional) point $\mathbf{q_o}$ from which the procedure will walk downhill reaching a minimum (probably a local one).

The algorithm starts with $\mathbf{q_o}$ defining the initial simplex (the other N points $\mathbf{q}_i$) as

$$\mathbf{q_i} = \mathbf{q_o} + \beta\varrho_i \tag{20}$$

where ϱ_i are the 3N unit vectors and β is a constant that is a guess for the length of the problem that can, in turn, have different values for its x,y,z components.

1. Further details can be found in the literature [15, 29].

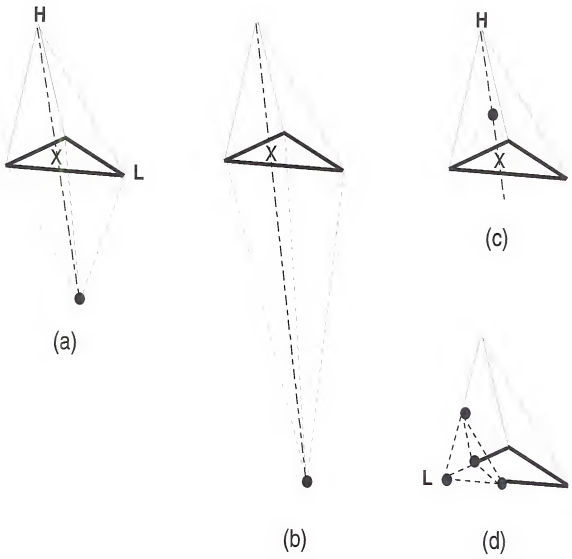


Figure 2.2. The Simplex method. The different steps that can be taken by the algorithm according to the topology of the potential energy surface. In all cases the original simplex is represented by solid lines. The generated simplex, represented (for all cases) by dashed lines, can be *(a)* A reflection in the opposite side of the triangle that contains the lowest energy point (L) to which the highest energy point (H) does not belong, *(b)* A reflection plus an expansion, *(c)* A contraction along the dimension represented by the highest energy point and the point where the triangle oppossite to it is crossed (X), or *(d)* A contraction along all dimensions.

The first steps are spent moving a projection of the highest energy point through the opposite face (to which it does not belong) that contains the point of lowest energy. This step is known as to be a reflection because it is taken in such a way that the volume of the initial simplex is held constant. Then the algorithm extends the new simplex in a given direction in order to take larger steps. When in the surroundings of a minima, the algorithm contracts itself in a transverse direction trying to softly spread down to the valley. The procedure can also, in these situations, contract itself in all directions in order to find a perhaps final tortuous minimum.

The search stops when the magnitude of the last displacement vector is fractionally smaller than a given threshold. In addition, it is customary to examine the decrease in energy such that simplex stops if this difference is smaller than a given threshold.

Restricted Step Method

Included here for historical and review reasons, Greenstadt has studied the Relative Efficiency of Gradient Methods up to the ones available until 1970 [23, 31].

This procedure rewrites the Taylor series expansion of the energy of Eqn. (3) as

$$E - E_o = \Delta E = \mathbf{q}^\dagger g + \frac{1}{2}\mathbf{q}^\dagger \mathbf{H}\mathbf{q} + \cdots \tag{21}$$

Now the following Lagrangian ($\mathcal{L}$) is introduced

$$\mathcal{L} = \Delta E - \frac{\lambda}{2}\left[\mathbf{q}^\dagger \mathbf{q} - h\right] \tag{22}$$

where $\mathbf{q}$ is the difference in coordinates (actual and previous cycle), λ is Lagrange's multiplier and h is the trust radius (or confidence region) for the stepping. The square bracket factor ensures that the search remains in the quadratic region.

The first derivative of the Lagrangian gives

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{0} = g + \mathbf{H}\mathbf{q} - \lambda \mathbf{q} \tag{23}$$

from here

$$\mathbf{q} = -\left[\mathbf{H} - \lambda \mathbf{I}\right]^{-1} g \tag{24}$$

where $\mathbf{I}$ is the identity matrix. This is a Newton-Raphson-like procedure which is fully recovered when $\lambda = 0$. By construction this technique has shown to be useful for transition state search as well.

Rational Functions (RFO)

Created by Banerjee and his coworkers [32], this is a procedure in which the energy is written in a normalized form as

$$\Delta E = E - E_o = \frac{\mathbf{q}^\dagger g + \frac{1}{2}\mathbf{q}^\dagger \mathbf{H} \mathbf{q}}{1 + \mathbf{q}^\dagger \mathbf{S} \mathbf{q}} \tag{25}$$

where $\mathbf{S}$ is a step matrix. Next step is to *augment* all the components of this rational function such that, and by using Eqn. (25) the difference in energy now becomes

$$\Delta E = \frac{\frac{1}{2}\left\langle \mathbf{q}^\dagger \quad 1 \right| \begin{vmatrix} \mathbf{H} & g \\ g & 0 \end{vmatrix} \left| \mathbf{q} \quad 1 \right\rangle}{\left\langle \mathbf{q}^\dagger \quad 1 \right| \begin{vmatrix} \mathbf{S} & g \\ g & 0 \end{vmatrix} \left| \mathbf{q} \quad 1 \right\rangle} \tag{26}$$

As usual, the next move is to ask for the first derivative of the change in energy with respect to the coordinates to be zero ($\partial \Delta E / \partial \mathbf{q} = 0$) to obtain the following eigenvalue equation

$$\begin{vmatrix} H & g \\ g^\dagger & 0 \end{vmatrix} \left| \mathbf{q} \quad 1 \right\rangle = \lambda \left| \mathbf{q} \quad 1 \right\rangle \tag{27}$$

from which two sets of equations are obtained

$$\begin{aligned} \mathbf{H}\mathbf{q} + g &= \lambda \mathbf{q} \\ g^\dagger \mathbf{q} &= \lambda \end{aligned} \tag{28}$$

The first of these two relations gives a Newton-Raphson-like step

$$\mathbf{q} = -\left(\mathbf{H} - \lambda \mathbf{I}\right)^{-1} g \tag{29}$$

which is, as before, recovered for $\lambda = 0$.

Reaction Path Following Method

This method requires one to know the TS, which will connect in a steepest descent fashion, the TS with reactants and products in order to draw a reaction path. First developed by Gonzalez and Schlegel [10], they have included modifications from the very beginning in order to consider the effects due to atomic masses. In subsequent articles they proposed a "Modified Implicit Trapezoid Method", which is a contribution on the way of obtaining the final coordinates used by this method, namely: "The Constrained Optimization" by Gear [33]. The method will be described in the TS section when we discuss Schlegel's procedure as it is related to the search of maxima and mimima.

The Hellmann-Feynman Theorem

The idea [34] is to obtain gradients to be used to optimize molecular geometries. To accomplish this, the starting point is Schroedinger's equation

$$|\mathrm{H}|\Psi\rangle = E_o|\Psi\rangle \tag{30}$$

or,

$$\langle\Psi|\mathrm{H}|\Psi\rangle = E_o\langle\Psi|\Psi\rangle = E_o \tag{31}$$

where the wave function is assumed to be normalized to unity $\langle\Psi|\Psi\rangle = 1$. The demonstration of the theorem starts by taking the first derivative of the energy with respect to a given (set of) coordinate(s) "q" as follows

$$\frac{\partial E_o}{\partial q} = \left\langle \frac{\partial\Psi}{\partial q} \middle| \mathrm{H} \middle| \Psi \right\rangle + \left\langle \Psi \middle| \mathrm{H} \middle| \frac{\partial\Psi}{\partial q} \right\rangle + \left\langle \Psi \middle| \frac{\partial \mathrm{H}}{\partial q} \middle| \Psi \right\rangle = 0 \tag{32}$$

If the two first terms vanish,

$$\left\langle \frac{\partial\Psi}{\partial q} \middle| \mathrm{H} \middle| \Psi \right\rangle + \left\langle \Psi \middle| \mathrm{H} \middle| \frac{\partial\Psi}{\partial q} \right\rangle = 0 \tag{33}$$

then the gradient will be simply

$$\frac{\partial E}{\partial q} = \left\langle \Psi \,\middle|\, \frac{\partial \mathrm{H}}{\partial q} \,\middle|\, \Psi \right\rangle \tag{34}$$

This is a very appealing relation since $\partial \mathrm{H}/\partial q$ is easily obtained. Experience tells us that this scheme works only for very, very good wave functions. It is obvious why since, for an exact eigenfunction, equation (33) can be written as

$$E_o \left[\frac{\partial}{\partial q} \left\langle \Psi \,\middle|\, \Psi \right\rangle \right] = 0 \tag{35}$$

This step concludes the proof because the bracket involves the first derivative of a constant, due to the normalization condition, and consequently the term vanishes.

Transition State Search Methods

Introduction

It is clear that procedures to locate TS and geometry optimization (GOPT) algorithms are intimately related. Techniques to find minima have been much more successful than those developed for locating TS. Their success is based on the relative ease in following downhill searches, such as with the steepest descent type of algorithms. Because of this success the problem of locating TS repeatedly has been approached as one of dealing with the location of minima. In general, such methods choose a higher energy point and from there walk downhill, stopping at a local minimum where the signature of the Hessian is checked (one and only one negative vibrational mode). But when studying a reaction mechanism, knowledge of the lowest energy reaction path is of great use but expensive. To accomplish this the recipe is to, by sitting at the found TS, follow a downhill path to reactants and products.

Other methods use a *mirror-image* technique. Consider the picture of reactants and products (usually only one of them) as minima and TS as a lowest energy maxima

between them. Placing a mirror at the TS the image obtained is the one of two maxima (reactants and products) and a minimum (original TS). Consequently the problem now is to find that minimum. Again downhill methods are used with the pitfalls described above.

A collection of the procedures reviewed here, showing their general features, advantages and disadvantages can be found in Table 2.1.

Simple Monte Carlo and Simulated Annealing Algorithms

The Metropolis Monte Carlo algorithm [35] has proven very successful in evaluating equilibrium properties of systems. The bulk properties are simulated from a small physically meaningful number of particles (N) such that the fluctuations in the calculated value of a property, usually a thermodynamic observable, are minimized.

The interactions of N particles are described by a potential energy function, say $U(r,\Omega)$, where **r** is the distance separating the particles while Ω represents any other coordinates (Eulerian angles for example) on which the potential may depend. Therefore, the potential interaction between particles ΔE is written as

$$\Delta \mathrm{E} = \sum_{i,j,..N} .. \sum_{r_i,r_j..N} .. \sum_{\Omega_i,\Omega_j..N} U(r_i, r_j, \Omega_i, \Omega_j, ...\mathrm{N}) \tag{36}$$

Monte Carlo method

N particles are placed in a system of volume **V** such that the macroscopic density is kept constant. The initial configuration of the particles is arbitrary, a flexibility which is a tremendous advantage of this method.

The position of any particle i with a position $r_i = (x_i, y_i, z_i)$ is chosen randomly and moved according to

$$X_i = x_i + bn_1 \qquad Y_i = y_i + bn_2$$

$$\vdots$$

$$Z_i = z_i + bn_3 \quad .... \quad \Omega_i^{'} = \Omega_i + fn_\kappa \tag{37}$$

where **b** and **f** are chosen step sizes for **r** and Ω and $\{n_i\}$ is a set of random numbers (for all $i \epsilon \mathbb{Z}$) in the interval [-1, 1]. The particle always stays in the cube such that surface effects are reduced and the density of particles (ρ) is always constant. The new conformation energy $\Delta \mathrm{E}^{'}$ is calculated according to Eqn. (36).

If $\Delta \mathrm{E}^{'} < \Delta \mathrm{E}$ then the new arrangement of the particle is accepted, and the calculation is continued from the newer configuration. On the other hand if $\Delta \mathrm{E}^{'} > \Delta \mathrm{E}$ then a probability (**P**) is calculated as

$$\mathrm{P} = exp\left[-(\Delta \mathrm{E}^{'} - \Delta \mathrm{E})/\mathrm{K}_B \mathrm{T}\right] \tag{38}$$

where K_B and T refer to Boltzman's constant and the temperature, respectively. Thus one of the following conditions is met:

If $\mathrm{P} > \delta$, $\delta \epsilon [0, 1]$ the move is accepted

If $\mathrm{P} < \delta$, $\delta \epsilon [0, 1]$ the move is rejected

In the first case the algorithm continues as described above. In the second case, where the new configuration is rejected, the particle is returned to its initial position and a new particle is chosen randomly and moved according to Eqn. (37) (δ is a random number).

The method is repeated until no further configurations are accepted. The system is said to have reached its equilibrium configuration when this criterion is satisfied.

The efficiency in reaching the minimum by Metropoli's algorithm depends on the number of moves allowed for the displacement of particles. A more accurate equilibrium

configuration of the particles can be determined if the number of random moves allowed is large. These methods are not competitive with gradient methods in obtaining local minima but, as discussed below, Monte Carlo allows us to leave a region of local minimum for the global one.

Simulated Annealing method

Simulated Annealing is similar to the Monte Carlo algorithm. The difference is that the probability **P** is evaluated as

$$\mathrm{P} = exp\left[-(\Delta \mathrm{E}' - \Delta \mathrm{E})/\mathrm{T}^*\right] \tag{39}$$

where T^* is a parameter with energy units. The potential energy surface is scanned in a finite number of moves using the random process described above for a given value of T^*. Then T^* is varied using an annealing factor α as follows

$$\mathrm{T}^*_{i+1} = \alpha \mathrm{T}^*_i \qquad , \qquad 0 < \alpha < 1 \tag{40}$$

where i is the number of steps allowed. The search process is then repeated with the new value of T^*.

As T^* decreases, areas of the surface closer to the minimum are scanned and if any minima had been missed by the search using the previous value of T^*, the method can now lock itself onto a lower minimum. As the number of cycles allowed for the annealing step increases, the search for the global minimum of lower energy becomes more efficient. This flexibility in being able to "anneal" the PES is one of the assets of the simulated annealing method.

Differing from gradient techniques, in which the displacements are generally within a small region of the PES, the random displacements in annealing enable the search to tunnel out of local minima in which the algorithm could have been trapped. The

evaluation of the global minimum can be assured provided that the algorithm has been allowed a large enough number of random moves for each value of T^* [35]. A discussion of these algorithms, with applications, can be found in the Allison and Tildesley's text [36].

Synchronous-Transit Methods (LST & QST)

Initially proposed by Halgren and Lipscomb [37], the Linear Synchronous Transit (LST) and the Quadratic Synchronous Transit (QST) methods treat "Forward" and "Reverse" processes equivalently, generating a continuous path between specified R and P. The main features of this technique are schematically shown in Figure 2.3.

The LST pathway is constructed by considering a linearly interpolated internuclear distance connecting reactants and products and estimating a TS that is improved by minimizing the energy with respect to all perpendicular coordinates. Finally the reaction path is approximated using a parabolic path between R and P, that is, the QST path giving a good estimation of the TS location.

The path coordinate (PC) (steps) is defined as: PC=DR/(DR+DP), where DR and DP are a measure of the distance to the path-limiting-structure, obtained as

$$DQ = \left[\frac{1}{N}\sum_{i=1}^{N}(X_i^m - X_i^q)^2 + (Y_i^m - Y^q)^2 + (Z_i^m - Z^q)^2\right]^{1/2} \qquad (41)$$

where Q and q = R,P (reactants or products), N is the number of atoms and m stands for, in accordance to the principle of least motion (PLM) [38], optimized structures of reactants and products that are re-oriented relative to each other in terms of rigid translations and rotations such that the sums over squares, for all corresponding atom coordinate differences (equation (41)), reaches a minima.

Intramolecular distances $R_{\alpha\beta}$ must vary simultaneously between the path-limiting structures $R^R_{\alpha\beta}$ and $R^P_{\alpha\beta}$. To avoid limitations in the method, provision must be taken to meet one of the following two conditions:

a) Linear Condition:

$$R^{(i)}_{\alpha\beta} = (1-f)R^R_{\alpha\beta} + fR^P_{\alpha\beta} \qquad 0 \le f \le 1\ ; \quad \alpha < \beta = 1, \mathrm{N} \tag{42}$$

b) Parabolic Condition:

$$R^{(i)}_{\alpha\beta}A = A + B\cdot f + C\cdot f^2 \qquad 0 \le f \le 1\ ; \quad \alpha < \beta = 1, \mathrm{N} \tag{43}$$

where

$$\begin{aligned} A &= R^R_{\alpha\beta} \\ B &= R^P_{\alpha\beta} - R^R_{\alpha\beta} - C \\ C &= \left[R^M_{\alpha\beta} - \big(1-\mathrm{PM}\big)R^R_{\alpha\beta} - \mathrm{PM}\cdot R^P_{\alpha\beta}\right] \Big/ \left[\mathrm{PM}\big(\mathrm{PM}-1\big)\right]. \end{aligned} \tag{44}$$

This ensures that the following conditions are satisfied

$$\begin{aligned} f = 0 \quad &\rightarrow \quad R^{(i)}_{\alpha\beta} = R^R_{\alpha\beta} \\ f = 1 \quad &\rightarrow \quad R^{(i)}_{\alpha\beta} = R^P_{\alpha\beta} \\ f = \mathrm{PM} \quad &\rightarrow \quad R^{(i)}_{\alpha\beta} = R^M_{\alpha\beta} \end{aligned} \tag{45}$$

where f is the interpolation parameter, i refers to interpolated quantities, PM denotes the value for the path coordinate and $\left\{R^M_{\alpha\beta}\ ;\ \alpha < \beta = 1, N\right\}$ are the atomic distances of some intermediate structure on the path. Geometries of the synchronous transit path may be evaluated in terms of interatomic distances from equations (42) and (43). In practice, linearly/parabolic interpolated Cartesian coordinates between path-limiting structures at maximum coincidence are subsequently refined so as to minimize

$$\sum_{\alpha=1}^{\mathrm{N}-1}\sum_{\beta=\alpha+1}^{\mathrm{N}}\left(\frac{1}{R^{(i)}_{\alpha\beta}}\right)^4\left[R^{(i)}_{\alpha\beta} - R^{(c)}_{\alpha\beta}\right]^2 + 10^{-6}\times\sum_{\omega=x,y,z}\sum_{\alpha=1}^{\mathrm{N}}\left[W^{(i)}_{\alpha} - W^{(c)}_{\alpha}\right]^2 \tag{46}$$

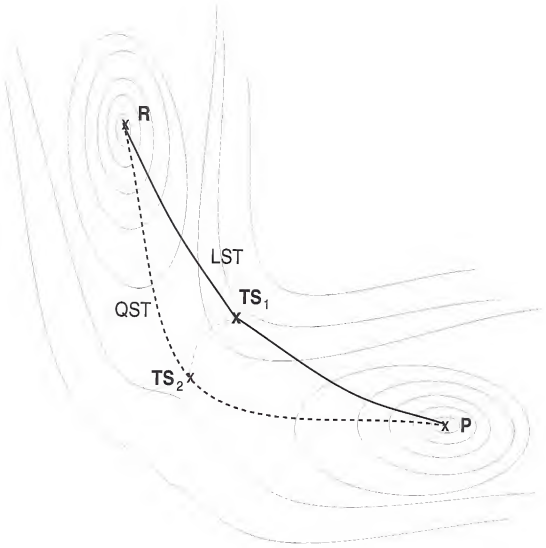


Figure 2.3. Potential energy surface representating Halgren-Lipscomb's TS search technique. The continuous line connecting reactants (R) and products (P), represents the LST with a maxima (TS) in TS_1. The dashed line represents the QST that passes through a transition state TS_2. The QST path has all the features required to represent the minimum energy reaction path. TS_1 and TS_2 are connected through a parabola. This issue has been discussed by Jensen through the Minimax/Minimi procedure [39].

where (i) stands for interpolated and (e) for evaluated (calculated) quantities referring to the evaluated (updated) Cartesian coordinates. The weighting factor $\left(1/R_{\alpha,\beta}^{(i)}\right)^4$ ensures a close reproduction of bond distances, whereas the 10^{-6} factor is proposed to suppress rigid translations and rotations, between the interpolated and calculated points ($W_\alpha^{(i)}$ and $W_\alpha^{(c)}$, respectively).

This procedure can be used for molecules with N>3 since the number of interatomic distances exceed the number of 3N-6 internal degrees of freedom for non linear molecules.

Cartesian coordinates are then submitted to the PLM to associate a unique path coordinate and the total energy is computed. A variation of f will produce a continuous energy path called, depending on the path, **LST** or **QST**. For example, in a uni-molecular reaction, the path will usually connect both limiting structures via some maximum path, whose structure can be determined using Eqns. (42) and (43).

Alternative algorithms have been developed maximizing along a path of known form and minimization perpendicular to the path [8, 25, 39]. In particular, Jensen has lately introduced a variation of this procedure namely the MINIMAX / MINIMI procedure [39], that is briefly discussed below.

Minimax / Minimi Method

Based on the Synchronous Transit method, the minimax/minimi method is a procedure for the location of transition states and stable intermediates [38]. It is based on the idea that a simple parabolic transit path cannot provide a correct description of the true minimum energy path, as suggested by Halgren and Lipscomb [37], if this path shows frequently changing sign of curvature.

An essential supposition states that to find a new quadratic synchronous transit maximum with higher energy after exhaustive orthogonal minimization is too expensive. Consequently, it is assumed that any practical method must explicitly take into account the influence of each geometric modification on the new transit path maximum.

It is then suggested that a straightforward way to proceed is the following: A change in a structure corresponding to the transit maximum (a minimum) under investigation will be accepted only if the resulting new path maximum (minimum) is of lower energy. Successive geometry optimizations (GOPT) of all internal coordinates will consequently lead to the lowest QST maximum, the transition state (TS), or to an intermediate (a local minimum) and is, because of this, called the MINIMAX / MINIMI optimization procedure [39].

A major drawback of this procedure is that an extra parabolic line minimization along the QST path, at each level of the parameter optimization, is needed. However, the procedure has the advantage that unexpected intermediates (MINIMAX) will be uncovered and that extreme shifts of the path coordinate may be obtained.

The Chain and Saddle Methods

The aim of this algorithm is to ensure stability towards a transition state. Stepping along the vector gradient field of an arbitrary continuous path between reactants and products, leads to a limiting path where the highest energetic point is considered the saddle point or, in the case of a multi-step mechanism, the highest energy transition state will be located. Figure 2.4 shows the behaviour of this technique.

The algorithm consists in replacing a chain of points $C(n) = (R, ..., p_i, ..., P)$ running from R to P by a new chain C(n+1) at iteration n+1. In order to maintain the connectivity of the path, each distance between two successive points is restricted to a

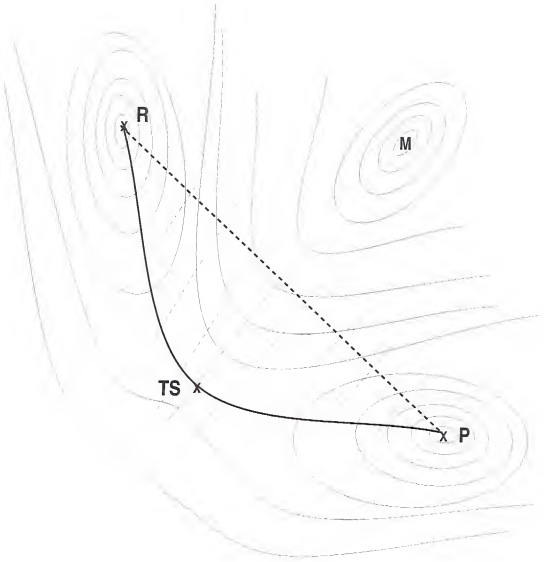


Figure 2.4. Potential energy surface representating the Saddle TS search technique. The dashed line connecting reactants (R) and products (P), represents the displacement vector from which identical fractions are taken as steps. At each step the energy is minimized in perpendicular directions (doted lines) to obtain new sets of projected coordinates that represents the minimum energy reaction path. The process is repeated until a maximum (TS) is found.

given length (in AMPAC [40] this length is 0.3 Å). The iteration consists of skipping the current highest point of the chain along either a descending or ascending path. In the first case the energetic relaxation of the whole path is insured, while on the second an interpolation of a point along the path is performed. New points are inserted as soon as a link length becomes too long. The successive evaluations of the gradient are used to update a quadratic local estimate of the potential, providing quadratic termination properties. This would make this procedure very computer time consuming.

Although this seems to be a Hessian free algorithm, it is not because a differentiation of a first order expansion of the energy is used [41]. Its recent appearance and lack of verification will exclude this procedure from Table 2.1. For extended references also see [24, 40–42].

Cerjan-Miller

This is essentially an uphill procedure [43] that is able to generate the reaction path coordinate by connecting a transition state with a minimum on the potential surface, schematically shown in Figure 2.5. It considers the Lagrangian function:

$$\mathcal{L}(\mathbf{s},\lambda) = E_o + \mathbf{s}^\dagger g + \frac{1}{2}\mathbf{s}^\dagger \mathbf{H}\mathbf{s} + \frac{\lambda}{2}\left(\Delta^2 - \mathbf{s}^\dagger \mathbf{s}\right) \tag{47}$$

where λ is the Lagrange multiplier, $\mathbf{s}$ is a fixed step size (the radius of the hypersurface), g is the gradient and $\mathbf{H}$ is the Hessian matrix. The extrema are determined by the conditions

$$\frac{\partial \mathcal{L}}{\partial \mathbf{s}} = 0 \qquad ; \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 0 \tag{48}$$

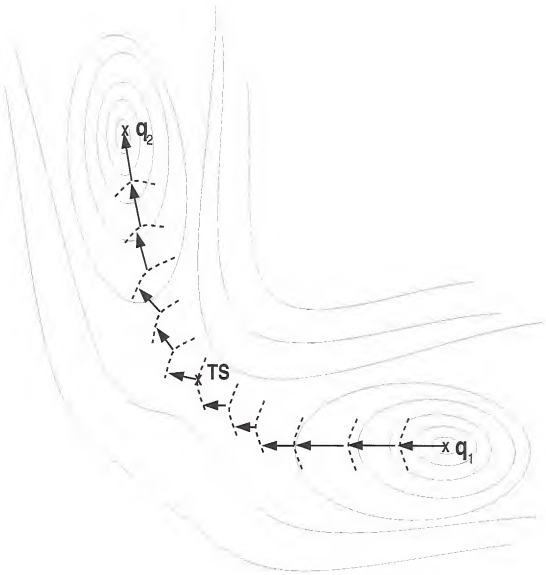


Figure 2.5. Potential energy surface representating Cerjan-Miller's technique. The arrows represent the step size Δ of the the trusted radius (dashed lines). The search starts from the minimum (q_1, reactants for example) and climbs up-hill towards the transition state (TS), from where a minimization is carried out to connect q_1 and TS with the new minima q_2 (products).

which gives the two following relations

$$\Delta^2 - \mathbf{s}^\dagger \mathbf{s} = 0 \qquad \rightarrow \qquad \Delta^2 = g^\dagger (\mathbf{H} - \lambda \mathbf{I})^{-2} g \tag{49}$$

where λ is evaluated for a given value of **s**. From Eqn. (49) the step size is obtained. E(s) is given then by

$$E(s) = E(\lambda) = E_o + g^\dagger (\mathbf{H} - \lambda \mathbf{I})^{-1} \left(\frac{1}{2}\mathbf{H} - \lambda \mathbf{I}\right) (\mathbf{H} - \lambda \mathbf{I})^{-1} g \,. \tag{50}$$

Now the unitary matrix **U** that diagonalizes **H**, is introduced: $\mathbf{U}^\dagger \mathbf{H} \mathbf{U} = \kappa$. At this point a new parameter "d" is defined: $d = \mathbf{U}^\dagger \cdot g$, then we write Eqn. (49) as

$$\Delta^2 = \Delta^2(\lambda) = \sum_{i=1}^{F} \frac{d_i^2}{(\lambda - \kappa_i)^2} \qquad : \qquad E(\lambda) - E_o = \sum_{i=1}^{F} \frac{d_i^2 (\lambda - \frac{1}{2}\kappa_i)}{(\lambda - \kappa_i)^2} \tag{51}$$

where the kappas (κ) are positive values for minima. Now the assumption that one is seated at a local minimum is made by saying that $\lambda = \lambda_o$, then it follows that $E(\lambda_o) - E_o > 0$, that is, the step **s** generated is indeed uphill in this direction. For λ = 0 the increment **s** (on Eqn. (49)) is the Newton Raphson step $\mathbf{s} = -\mathbf{H}^{-1} g$.

Finally the step for walking uphill to a transition state from the minimum of the potential surface is given by equation (51) (where $\lambda = \lambda_o$, if $\lambda_o > 0$). It must be noted that λ_o is the local minimum of the function $\Delta^2(\lambda)$, in other words, it is the root of

$$\frac{d\Delta^2(\lambda)}{d\lambda} = -2 \sum_{i=1}^{F} \frac{d_i^2}{(\lambda - \kappa_i)^3} = 0 \,. \tag{52}$$

In a general case, the function $\Delta^2(\lambda)$ will have F-1 local minima. Cerjan and Miller suggest picking up the smallest value of λ, that is, the smallest root of equation (52), corresponding to the softest mode.

Draw backs of this procedure are the use of second derivatives, the use of too many steps when approaching the transition state and the coupling between the step and the curvature radii of the surface in the actual point. This last issue is important as the next step might not encounter a minimum.

Schlegel's Algorithm

This is essentially a gradient algorithm [44], proposed in 1982 by Bernhard Schlegel[2], in which the "right inertia of the approximate Hessian matrix" is obtained by " adjusting the sign of inadequate eigenvalues." The sign of the smallest positive eigenvalue is changed if in the search of the TS no negative eigenvalue is present. On the other hand, if sundry negative eigenvalues happen, all of them are replaced by their absolute value (except the smallest one). Given the stationary condition $\nabla q^k(\mathbf{s}) = 0$ the quasi-Newton step, at cycle *k* in the step direction (**s**), is thus modified according to

$$s^k = -\sum_{i=1}^{n} \frac{\bar{g}_i^k}{b_i'} V_i^k \qquad ; \qquad |b_i'| = |b_i^k|, b_1' < 0 < b_2' < \cdots < b_n' \tag{53}$$

where the b_i' terms are eigenvalues of the Hessian $\mathbf{H}^k$, V is the eigenvector basis, g^k is the gradient and *k* is the cycles index. This is then related to Greenstadt's proposal[3], that is used in a minimization process. His idea is to reverse the ascendant/descendant character of the search direction. Nevertheless, in areas of large curvature, the resulting direction is not necessarily the opposite of the initial one, if the investigated region is far from an extremum and thus may be incorrect. A scaling factor is used to modulate this effect.

If the quasi-Newton search direction of Eqn. (53) exceeds the maximum allowed step R_{max}, its length is set to this maximum value. This change requires the addition of a shift parameter λ obtained by the search of an extremum of the quadratic function $q^k(g)$. In practice, the shift parameter λ is obtained by minimizing the function

$$\left(\| s^k(\lambda) \| - R_{max} \right)^2 . \tag{54}$$

2. We will keep here the super **k** indices used by Schlegel, according to Powell's notation; see: M.J.D. Powell, Math. Prog. **1:26** (1971).

3. For details see J. Greenstadt, Math. Comp. **21**, 360 (1967) ; Y. Bard *Nonlinear Parameter Estimation*. Academic Press, New York, 1974, p. 91–94.

The radius R_{max} is updated using a trust region method. The step direction is

$$s^k = -\sum_{i=1}^{n} \frac{\bar{g}_i^k}{(b_i' - \lambda)} V_i^k \,. \tag{55}$$

When implementing trust region methods, the minimization of Eqn. (54) is performed by determining the zero of its first derivative using a Newton-Raphson procedure. However, the convergence threshold of such an algorithm is guided by a zero value of Eqn. (54). Given that the minimum of this function is not necessarily associated with a zero function value, the procedure may fail.

Besides, this Newton-Raphson search of a zero value of the first derivative implies that the parameter λ lies in the open interval $]b_1,b_2[$. Thus, concerning Eqn. (55), the step s^k is uphill along the first eigenvector V_1^k and down-hill along all the others.

The Normalization Technique or E Minimization

Developed by Dewar and co–workers [41, 42], the Normalization Technique is a root search technique rather than a saddle point location. Only convergence to a zero of the gradient is ensured, not necessarily the TS. Moreover, the procedure has been shown to require a good initial guess. In fact, if the PES is tortuous, stability problems appears and the procedure requires a large number of energy evaluations to be successful.

Originally implemented in the closed-shell version of MNDO, the geometry of reactants (R) and products (P) is defined (in 3N–6 coordinates) as $R = \sum_i a_i$ and $P = \sum_i b_i$. A reaction coordinate (D) is defined as:

$$R - P = D = \sqrt{(a_i - b_i)^2} \tag{56}$$

where D is reduced subject to the condition that the structure with lower energy is moved to approach the TS. The following procedure is used:

1) Obtain the optimized geometry of R and P.

2) Evaluate the energy of R and P then, defining the origin on the higher energy structure, the geometry of the other species is expressed in terms of its new origin as:

$$\sum_i a_i' = \sum_i (a_i - b_i) \qquad \rightarrow \qquad D = \sum_i (a_i')^2 \tag{57}$$

3) Modify geometry of lower energy structures to select a new distance[4] D' to reduce the difference between R and P as: $a_i = \sum a_i' D'/D$.

4) Optimize R's geometry such that D is held constant at D'.

5) If D is small enough, then stop; otherwise go to step 3.

Caution must be taken in ensure that one geometry (for example products) can be obtained from the other (for example reactants) by a continuous deformation [16, 17, 42, 44–46]. The first work of Komornicki and McIver [46] is also known as the Normalization Energy Minimization or as the Gradient Norm Minimization technique. Pertinent previous work of Komornicki and McIver is cited in their last 2 articles in the literature [8, 16, 47].

Augmented Hessian

The Augmented Hessian procedure was originally proposed by Lengsfield [19, 48] for MCSCF calculations, and further developed by Nguyen and Case [49] and later on by many other groups [33, 50]. Augmented Hessian is essentially an uphill walking algorithm, implemented in the ZINDO package by Zerner and co-workers [19, 26]. The search direction **s** is found by diagonalizing the "Augmented Hessian"

$$\begin{pmatrix} \mathrm{H} & \alpha g \\ \alpha g^{\dagger} & 0 \end{pmatrix} \begin{pmatrix} \nu \\ \beta \end{pmatrix} = \lambda \begin{pmatrix} \nu \\ \beta \end{pmatrix} . \tag{58}$$

4. Typically D' = 0.95•D

For a down hill search, λ is the lowest eigenvalue, α a parameter that can be varied to give the required step lengths, and the lowest normalized eigenvector: $\nu^2 + \beta^2 = 1$. From Eqn. (58) we get two equations

$$\begin{aligned} \mathrm{H}\nu + \alpha\beta g = \lambda\nu \qquad &\rightarrow \qquad \begin{cases} (\mathrm{H} - \lambda)\nu = -\alpha\beta g \\ \nu = -(\mathrm{H} - \lambda)^{-1}\alpha\beta g \end{cases} \\ \alpha g^{\dagger}\nu = \lambda\beta \end{aligned} \tag{59}$$

solving for the step size we get

$$s = -(\mathrm{H} - \lambda\mathrm{I})^{-1}g = \nu/\alpha\beta \tag{60}$$

where **H** is the exact Hessian matrix or, as suggested by Zerner *et al.* [19], an approximate matrix of the Hessian if **H** is not available. The step direction is obtained by writing **s** in terms of the gradient (**g**) the eigenvalues and eigenvectors of the Hessian (λ_i and $|v_i\rangle$ respectively):

$$s = -\sum \frac{|v_i\rangle\langle v_i|g\rangle}{\lambda_i - \lambda} \,. \tag{61}$$

To ensure an uphill search direction, a specific eigenvector of **H** that overlaps strongly with the uphill search direction, is chosen such that λ_x is scaled using scaling factors and the search direction **s** is obtained as

$$s = -\frac{|nv_x\rangle\langle nv_x|g\rangle}{\lambda_x^{'} - \lambda} - \sum_i \frac{|v_i\rangle\langle v_i|g\rangle}{\lambda_i - \lambda} \tag{62}$$

where the scaling factor n is chosen as: $n = \sqrt{\lambda_x/\lambda_x^{'}}$ and $\lambda_x^{'} = \lambda_2/4$; λ is chosen to lie between $\lambda_1 < \lambda < \lambda_2$. Thus the step is scaled accordingly to the curvature of the quadratic region. Finally, the stationary point χ_e is found switching to the Norm of the Gradient Square Method (**NGSM**, see below) when **H** develops a negative eigenvalue [19].

Although this model has the advantage of being precise, it is expensive to compute since the exact Hessian is required. It has to be pointed out that Jensen and Jorgensen

[51] developed this method for MCSCF optimization of excited states. Further developments were carried out by Zerner and his co-workers [27].

Norm of the Gradient Square Method (NGSM)

The sum of the squares of the gradient (g), written as

$$\sigma = \sum_i g_i^2 = \langle g|g\rangle \tag{63}$$

is minimized [18] as was initially suggested by McIver and Komornicki [16]. The Taylor series expansion will be:

$$\sigma_{\kappa+1} = \sigma_\kappa + \sigma_\kappa^{'}\mathbf{s}_\kappa + \frac{1}{2}\mathbf{s}_\mathbf{k}^{\dagger}\sigma_\kappa^{''}\mathbf{s}_\kappa + \ldots \tag{64}$$

and

$$\sigma_{\kappa+1}^{'} = \sigma_\kappa^{'} + \sigma_\kappa^{''}\mathbf{s}_\kappa \tag{65}$$

where k indexes the cycles and s_κ is the step, defined as·

$$s_\kappa = \mathrm{X}_{\kappa+1} - \mathrm{X}_\kappa \tag{66}$$

From Eqn. (65), an extreme point for the function σ is one in which $\sigma_{\kappa+1}^{'} = 0$ then

$$s_\kappa = -\left(\sigma_\kappa^{''}\right)^{-1}\sigma_\kappa^{'} \tag{67}$$

where

$$\begin{aligned}\sigma_{jk}^{'} &= \frac{\partial\sigma}{\partial X_j} = 2\sum_i g_i\frac{\partial g_i}{\partial X_j} \\ \sigma_{jk}^{''} &= \frac{\partial^2\sigma}{\partial X_j\partial X_k} = 2\sum_i\left(g_i\frac{\partial^2 g_i}{\partial X_j\partial X_k} + \frac{\partial\sigma}{\partial X_j}\frac{\partial g_i}{\partial X_k}\right)\end{aligned} \tag{68}$$

or in matrix form

$$\sigma^{'} = 2\mathbf{H}g \qquad\qquad \sigma^{''} = 2\left[\mathrm{C} + \mathbf{HH}\right] \tag{69}$$

It must be noticed that here

$$C_{jk} = \sum_i g_i \frac{\partial^2 g_i}{\partial X_j \partial X_k} \tag{70}$$

contains the third derivative $\partial^3 E/\partial X_i \partial X_j \partial X_k$ and becomes less important as $g_i \rightarrow 0$, that is, as an extreme point is approached.

From Eqn. (67) we have

$$s = -\left[2(\mathbf{C} + \mathbf{HH})\right]^{-1} 2\mathbf{H}g = -\left[\mathbf{C} + \mathbf{HH}\right]^{-1} \mathbf{H}g \,. \tag{71}$$

When $\mathbf{C} \rightarrow 0$:

$$s = -\mathbf{H}^{-1} g \tag{72}$$

which is the Newton-Raphson equation for an extremum point provided that $\mathbf{C}$ is sufficiently small (locally, near an extreme point it must always be correct).

The "object" function being reduced from equation (64) is σ (not E) and the line search condition is: $\sigma_{\kappa+1} < \sigma_{\kappa}$.

This method can be applied to find any stationary point and will not necessarily find local minima with respect to the energy: Rather, one usually increases the energy of the nearest stationary point and then minimize it with this technique.

Gradient Extremal

This model, first proposed by Ruedenberg [52] and further developed by others [53], uses gradient extremals which are defined as lines on the mass scaled potential energy surface E(x) having the property that, at each point x_o, its molecular gradient g(x_o) is a minimum with respect to variations within the contour subspace, for example, along a contour of E(x) constant. Figure 2.6 shows the behaviour of this procedure as it steps uphill, whereas Figure 2.7 shows how minima and maxima get connected through the gradient extremal.

The procedure starts by introducing the Lagrangian multiplier λ

$$\partial\left[g^{\dagger}g - 2\lambda(E-\kappa)\right]/\partial x = 0\ . \tag{73}$$

By differentiating Eqn. (73) the following eigenvalue equation is obtained

$$\mathbf{H}(x)g(x) = \lambda(x)g(x)\ . \tag{74}$$

This is perhaps the most important contribution of this technique, as it states that the gradient is an eigenvector of the Hessian. A simple interpretation of this expression is that $2\mathbf{H}\mathsf{g}$ is proportional to the gradient g at the point $\boldsymbol{x}$. Moreover, g is orthogonal to the contour subspace at gradient extremals, since g is orthogonal to the contour subspace.

It is assumed that the potential energy, its gradient and Hessian are calculated explicitly at each iteration. Setting the geometry of the k'th iteration, say $\mathrm{x_k}$, a step $\mathrm{s_k}$ is determined, where it is possible to write: $x_{k+1} = x_k + s_k$. The second order total energy at this point is approximated as

$$\mathrm{E}^{(2)}(x_{k+1}) = E(x_k) + g^T(x_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}^{\dagger}_k\mathbf{H}(x_k)\mathbf{s}_k \tag{75}$$

and the actual energy at this point (with no approximations) is

$$\mathrm{E}(x_{k+1}) = \mathrm{E}^{(2)}(x_{k+1}) + R \tag{76}$$

where R contains higher order terms in $\mathbf{s_k}$. Steps are taken with confidence if: $\mathrm{E}^{(2)}(x_{k+1}) \rightarrow \mathrm{E}(x_{k+1})$. A quantitative measure of this approach to agreement may be obtained from the ratio r as

$$r = \frac{\left[\mathrm{E}(x_{k+1}) - \mathrm{E}(x_k)\right]}{\left[\mathrm{E}^{(2)}(x_{k+1}) - \mathrm{E}(x_k)\right]} = 1 + \frac{R}{\left[\mathrm{E}^{(2)}(x_{k+1}) - \mathrm{E}(x_k)\right]} \tag{77}$$

If $\mathrm{r}\rightarrow 1$, the third-order terms are negligible and the second-order expansion is considered to be exact. The chosen step size should then depend on how close r is to unity.

A trust region with radius h is introduced, within which the second-order expansion approximates the exact potential surface, and the trust radius is updated according to the size of r.

The step direction ($\mathbf{s}_k$) is obtained by using the extremal of the second-order surface. The steps in the walk are determined assuming that in the trust region the gradient extremal of the second-order surface will describe accurately the gradient extremal of the exact surface. In the quadratic region we have

$$\mathrm{H}(x) = \mathbf{H} \quad ; \quad \lambda(x) = \lambda \qquad \rightarrow \qquad g(x) = g + \mathbf{H}x \tag{78}$$

where H and λ are constant. It is assumed that the origin is the center of expansion. Substitution of Eqn. (78) in (74) gives

$$(\mathbf{H} - \lambda\mathrm{I})\mathbf{H}x = -(\mathbf{H} - \lambda\mathrm{I})g \tag{79}$$

which reproduces the Newton-Raphson step equation if $(\mathbf{H} - \lambda\mathrm{I})^{-1}$ exists. Let v be the eigenvector of H belonging to λ (the eigenvector along the reaction path): ($\mathbf{H}$-$\lambda\mathbf{I}$)v=**0**. If λ is non-degenerate then ($\mathbf{H}$-$\lambda\mathbf{I}$) is non-singular on the orthogonal complement of v. Thus, the following projector is introduced: $\mathbf{P} = \mathrm{I} - vv^{\dagger}$ and Eqn. (79) is now written as

$$\mathbf{P}\mathbf{H}x = -\mathbf{P}g \qquad \rightarrow \qquad \mathbf{P}x = -\mathbf{P}\mathbf{H}^{-1}g \tag{80}$$

The solution for this relation, assuming that $\mathbf{H}$ is non-singular, is

$$x(\alpha) = -\mathbf{P}\mathbf{H}^{-1}g + \alpha v \,. \tag{81}$$

The gradient extremal x(α) (alpha is an arbitrary real parameter) for the second-order surface defines a straight line which is parallel to the eigenvector v passing through the solution of the projected Newton equation $\mathbf{PH^{-1}g}$ where αv is the step in our Newton-Raphson scheme. If now the Hessian has the desired number of negative eigenvalues

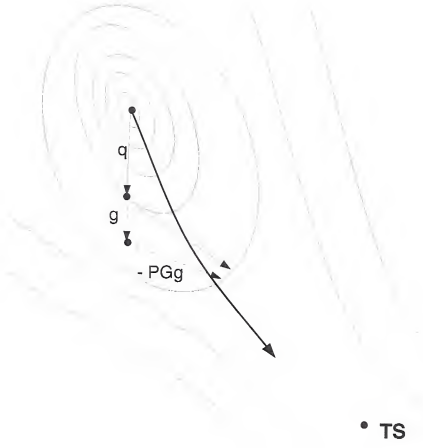


Figure 2.6. Potential energy surface representing the gradient extremal uphill walk (bold arrow) that will connect stationary points, that is, all minima and transition states.

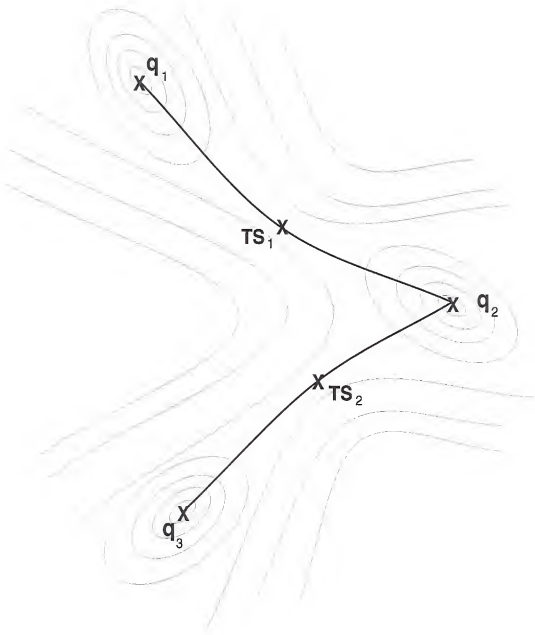


Figure 2.7. Potential energy surface representing gradient extremal, unique lines (bold) connecting stationary points, that is, all minima ($\mathbf{q_1}$, $\mathbf{q_2}$ and $\mathbf{q_3}$) and transition states (TS_1 and TS_2).

(only one for a true TS), then the stationary point of the surface is used as the next iteration point x_{k+1}. On the other hand, if the stationary point is outside the trust region or if the Hessian has not the desired index, then the gradient-extremal point on the boundary of what becomes the next iteration will point downhill. The gradient-extremal point on the boundary is determined by varying α in Eqn. (81) to obtain a step length equal to the current trust radio h.

Although results are promising for this procedure, **H** has the specific drawback that the step must be inside or on the boundary of the trust region and those steps are conservative. The gradient extremal has been found to bifurcate also during such a walk. It is important to note that the usefulness of the gradient extremal is related to the fact that there are unique lines connecting stationary points, as shown in Figure 2.7. This, together with the fact that these lines are locally characterized, makes gradient extremals potentially very useful for exploring potential energy surfaces and for some uses in molecular dynamics. Unfortunately, applications of this technique have not been reported yet.

Gradient Extremal Paths (GEP)

The original idea of Gradient Extremal Paths is due to J. Pancir [54] with subsequent testing by Muller [55]. A formal mathematical definition was given by Basilevsky [56]. Hoffman et al. [52] discussed the nature of GEPs with emphasis on their usefulness in molecular dynamics. They showed that third-order derivatives are very important to characterize GEPs. Jorgensen et al. [31d] were among the first in developing algorithms to find TSs in chemical reactions using second order GEP. Recent developments and applications have appeared for GEP [10, 12, 13, 62a]. Use of GEP to obtain molecular vibrations, as well as a good review of this model have been discussed by Almlof [57].

Constrained Internal Coordinates

Internal Coordinates [19, 58] are often preferred over Cartesians because they allow valence bond parameters (bond lengths, bond angles) to be constrained in a physically meaningful way as the remaining structure parameters are optimized.

Such procedures can be summarized in accordance to the following three steps:

- Series of minimizations constraining some coordinates.
- TS is the E_{max} with respect to the unconstrained coordinate(s).
- Energy is minimized with respect to all other coordinates.

An advantage of this procedure is that the Hessian is not required to reach the saddle point. A major drawback is that an important reaction coordinate must be identified in advance.

The Image Potential Intrinsic Reaction Coordinate (IPIRC)

Designed by Sun and Ruedenberg [12d], IPIRC is a transformation of Fukui's Intrinsic Reaction Coordinates [59, 60] transition state search procedure converted into an algorithm that searches for minima. IRC was originally proposed by Fukui [59] and later developed by others [60]. Andres et al. [61] applied the IRC to the addition reaction of CO_2 to $CH_3NHCONH_2$ using different semiempirical methods and Ab-Initio basis sets.

The strategy of this technique is as follows:

1) Diagonalize the inverse of the Hessian matrix: $C^{\dagger}H^{-1}C = \Lambda$.

2) Organize the eigenvalues of the diagonal matrix Λ in decreasing order:

$$\lambda_1 > \lambda_2 > \cdots\cdots \lambda_n.$$

3) Change the sign of the smallest eigenvalue λ_n .

4) Undiagonalize Λ and procede to minimize using a steepest descent procedure.

As a consequence, the transition state structure now becomes a minima (to be sought) and the original minima (starting conformation) becomes a higher energy structure from which the down-hill walk (minimization) will start.

The Constrained Optimization Technique

Constructed by Muller and Brown [62], the constrained optimization technique optimizes consecutively the geometry through a given pre-established coordinate. Abashkin and his collaborators as well as others [41, 63, 64], have proposed a mixture of techniques and implement this idea into DFT calculations. The main contribution of their algorithm is that they solve the problem of the constrained optimization by explicitly eliminating one of the variables using the constraint condition.

Gradient-Only Algorithms

A gradient-only algorithm recently was explored by Quapp [13]. It has as a major drawback its apparent necessity of a large number of steps to find the saddle point. Its success relies on the small size of the step it takes but, as a consequence, convergence is very slow. Figure 2.7 shows the main features of this procedure.

The algorithm starts by stating a new definition of the valley pathway: A point $\mathbf{q}$ belongs to a γ—minimum energy path (γMEP) if the gradient condition $g(\mathbf{q}) = g(\mathbf{q}_\gamma)$ holds, and is used to compare differences of gradient vectors. The new coordinates are given by: $\mathbf{q}_\gamma = \mathbf{q} + \gamma g(\mathbf{q})$. We immediately recognize the steepest descent like relation to obtain the new coordinates in this uphill walk (where the Hessian has been replaced by the identity matrix). Here γ is a step length parameter (not coming from a line search).

An asymptotic steepest descent path is defined as the geometrical space in which many steepest descent lines, from the left and the right side, converge into the stream

bed of the valley ground whose shape will be followed by the γMEP. The points close to these path are shown in Figure 2.2.

The situations to be encountered are as follows: If the point $\mathbf{q}_{0l}$ is at the *left* of the γMEP, then the negative gardient of $\mathbf{q}_{1l}$ will point it back to the *right*. Conversely, if the point $\mathbf{q}_{0r}$ is displaced to the *right* then the negative value of the gradient at $\mathbf{q}_{1r}$ will point to the *left*. The idea is that this gradients can be used to correct the steps as they go apart from γMEP.

The algorithm, which needs a step length (s) and a tolerance (t) to start, is as follows:

1) Optimize starting geometry $\mathbf{q}_0$ that it is not necessarily a minimum: $|g(\mathbf{q}_0)| \neq 0$.

2) Choose a step length and a tolerance (t) such that: Set counter $i = 0$ and $t << s$ with $t < 1$.

3) Predict a step in a steepest descent fashion: $\mathbf{q}_{i+1} = \mathbf{q}_i + \gamma g(\mathbf{q}_i)$.

4) If $|g(\mathbf{q}_{i+1})| < T$ then STOP, meaning that a saddle point has been located (T is a given threshold).

5) Get a scaling factor (f) (in braket notation): $f = \langle g(\mathbf{q}_{i+1}) | g(\mathbf{q}_i) \rangle$.

Here a backwards checking is performed:

If $f \geq 1 - t$ then: set $i = i + 1$, Go To step 3). Else:

6) Correction to the step: $\mathbf{q}_{i+1} = \mathbf{q}_{i+1} - \gamma g(\mathbf{q}_{i+1})$, set $i = i + 1$ and Go To step 3).

The technique seems to work well if $t < 10^{-2}$. This procedure is not competitive, for example, with the Approximate LTP technique of Cardenas-Lailhacar and Zerner [14], which requires one-fourth as many energy evaluations to get the same results.

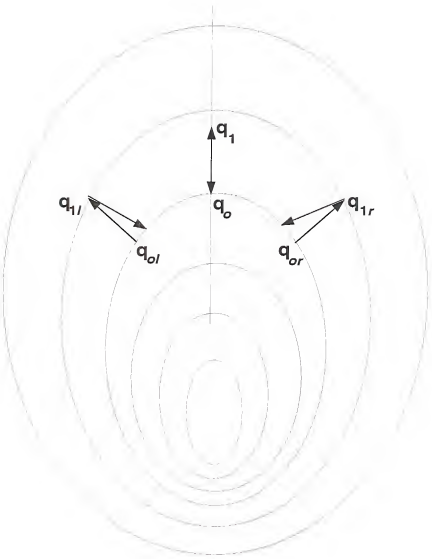


Figure 2.7. Quapp's only gradient procedure. Three different points on and in the neighborhood of a minimum energy path (MEP) are shown. The gradients (uphill arrows) are shown for points $\mathbf{q_{ol}}$, $\mathbf{q_o}$ and $\mathbf{q_{or}}$. For points $\mathbf{q_{1l}}$, $\mathbf{q_1}$, and $\mathbf{q_{1r}}$ the negative gradients (down hill arrows) are drawn. The uphill steps are corrected using the gradient vectors $-g(\mathbf{q_{1l}})$ and $-g(\mathbf{q_{1r}})$.

Table 2.1. Features, Advantages and Disadvantages of some of the most used Transition State search techniques available today in program packages.

	Model	Advantages	Disadvantages
Simple Monte Carlo	Partics. in volume. Arbitrary initial Config.: ρ = cte	Initial Config of the system is arbitrary	Random walk needs large number of moves
Simulated Annealing	Reduced Temperature to evaluate probability. ρ = cte	Flexibility to anneal the P.E.S.	Need large number of random moves
Synchronous Transit Path (LST and QST)	LST: Line connects R and P. QST: Max LST fitted to a Parabola	Simple assumptions about reaction path simplifies the search	If path is curved QST might not converge
Cerjan-Miller	Evaluate Hessian to define uphill path. Lagrange multiplier is used	Walks up-hill from minimum to TS essentially in an automatic way	Frequent H matrix calculation makes it expensive
Schlegel	Right inertia of App. H obtained by fixing eigenvalues sign	Reverse up/down search direction, refined by a factor	Fails if number of iterations needed is large. Downhill step - 1 dimension.
Minimax / Minimi	Successive Opt of Int. Coords. of a given Symmetry	Hints unexpected Intermediates or extreme shifts	Extra parabolic line minimization along QST
Energy Min or Normalization	Distance between R and P is used	Very simple and cheap procedure	Needs good initial guess for TS
Augmented Hessian	Search Dir founded diagonalizing the Approx. Aug. H	Precise, few cycles needed to Minimize the gradient	Evaluation of the H matrix is expensive
Gradient Extremal	Stationary points in PES connected by stream beds	Unique lines (**g**) that connect stationary points	Complications happen if Gradient Extremal bifurcates
Constrained Int. Coordinates	Selection of RC (bond length)	**g** is not needed to reach TS	Identify suitable Reac. Coord.
Squared Norm of the Gradient	Newton step-like search direction	Nearest stationary point uncovered	Costly evaluation of the Hessian matrix

CHAPTER 3
HISTORY OF GEOMETRY OPTIMIZATION AND TRANSITION STATE SEARCH IN SEMI-EMPIRICAL AND AB-INITIO PROGRAM PACKAGES

The relations describing the approximations involved in each model will not be examined here because this is not a comprehensive review. We refer the reader to the original works [65].

Brief Historical Overview

Semi-empirical molecular orbital theories are mainly based on approximations to the Hartree-Fock equations. The first of the Zero Differential Overlap (ZDO) methods that has historical importance is the π—electron method developed in 1931 by Hückel [66]. It is still used today to demonstrate important qualitative features of delocalized systems. In the early 1950's, Pariser, Parr, and Pople developed the PPP theory [67], which while only of historical importance, had great influence on future procedures. This technique was the first to describe molecular electronic spectroscopy with any degree of accuracy and generality. Related to these procedures, in 1952 Dewar developed the Perturbational Molecular Orbital (PMO) theory [68] (also a π electron method), calibrated directly on the energies of model organic compounds. The accuracy of this method was remarkable [69].

Pople and his co-workers, in 1965, extended the ZDO method to all valence electrons [70]. The impact that such an approximation had in the formation of the Fock matrix gave rise to new methods such as the Complete Neglect of Differential Overlap (CNDO), Intermediate Neglect of Differential Overlap (INDO), Neglect of Differential

Diatomic Overlap (NDDO) schemes. New modifications and approximations introduced in these procedures, have produced revised methods such as CNDO/1, CNDO/2, CNDO/S, INDO/1, INDO/2, and INDO/S.[5]The first gradient method was introduced by Komornicki and McIver [2] in CNDO and this important work had enormous impact.

The MINDO/3 model [71] (a Modified INDO model) was developed by Dewar and his collaborators in 1995. This technique was designed to reproduce experimental properties such as molecular geometries, heats of formation, dipole moments and ionization potentials. This method has prove to be remarkably useful. Introduced much later, MINDO/3 has an automatic geometry optimization procedure which was a contribution of tremendous impact at that time — the Davidon, Fletcher and Powell (DFP) [22] algorithm.

The SINDO1 model by Jug et al. has proven to be very accurate in reproducing geometries as well as other properties as binding energies, ionization potentials, and dipole moments [72]. The geometry optimization part of SINDO1 was implemented at the Quantum Theory Project (University of Florida) by Hans Peter Schluff, and uses the BFGS procedure [21, 27] as developed by Head and Zerner [19].

NDDO and MNDO

Proposed by M. Dewar and W. Thiel in 1977 [73], the Modified Neglect of Differential Overlap (MNDO) model was introduced as the first NDDO method. Today it uses the gradient norm minimization procedure for finding the transition state, whereas for optimizing geometries it employs a variation of the DFP algorithm [22].

MNDO, as well as the majority of other ab-initio and semi-empirical programs, is subject to improvements which generate a proliferation of related programs and methods

5. S stands for Spectroscopy; parameter sets are modified to reproduce electronic spectra.

such as MNDOC [74], initially parametrized only for H, C, N and O [75]. More recently, it has been suggested by M. Kolb and W. Thiel [76] that an improvement to the MNDO model can be achieved by the explicit inclusion of valence shell orthogonalization corrections, penetration integrals, and effective core potentials (ECP's) in the one-center part of the core Hamiltonian matrix. Their results shows good improvement in the location of TS over such methods as MNDO, AM1 and PM3. MNDO originally was parametrized on experimental molecular geometries, heats of formation, dipole moments, and ionization potentials.

MOPAC

In 1983 Stewart [77] wrote a semi-empirical molecular orbital program (MOPAC) containing both MINDO/3 and MNDO models, allowing geometry optimization and TS location using a Reaction Coordinate gradient minimization procedure, introduced by Komornicki and McIver [8] and vibrational frequency calculations.

AM1

The Austin Model 1 (AM1) developed in 1985 by Dewar and his group [41], was created as a consequence of shortcomings in the MNDO model (spurious interatomic repulsions, inability to reproduce hydrogen bonding accurately). Minima and TS location are the same as for MNDO.

PM3

The Parametric Method Number 3 (PM3) introduced by Stewart [78], is the third parametrization of the original MNDO model. As in AM1, PM3 is also a NDDO method using a modified core-core repulsion term that we will not describe here. PM3 and AM1 differ from each other in that PM3 treats the one-center, two-electron integrals as pure parameters. This choice implies that in PM3 all quantities that enter the Fock

matrix and the total energy expression have been treated as pure parameters. This, in turn, is proving to be a disadvantage as many anomalies are beginning to appear as a consequence of parameters that are not physically reasonable. Geometry is searched using the Saddle technique [40, 41].

ZINDO

This package of programs implemented by Zerner and co-workers [26], contains the INDO/1 and INDO/S models. ZINDO is constructed to perform a series of calculations based on different models, namely PPP, EHT, IEHT, CNDO/1, CNDO/2, INDO/1, INDO/2 and MNDO. It includes techniques to examine geometry [17, 19, 79] using the Line Search (see chapter 2), Newton-Raphson, Augmented Hessian, Minimize Norm Square of Gradient, and other techniques. This allows the user to select from a variety of search types as well as updating procedures (e.g. BFGS, Murtagh-Sargent (MS), DFP, and Greenstadt). For TS structures search, the "Augmented Hessian" procedure developed by Nguyen and Case [49] has been implemented by Zerner et al. [19]. Although this is a very effective method, it requires the use of second derivatives, again making it very time consuming. The Gradient Extremal method of Ruedenberg et. al. [52] is very effective, but requires the exact evaluation of the Hessian. The LTP algorithm [14] has recently been implemented; it makes use of up-date techniques like BFGS [21] to treat the second derivative matrix, although it uses reactants and products for the search. Widely used, the BFGS update procedure for geometry optimization and TS search was born in this versatile package; and is now used in Gaussian, HONDO, Gamess as well as AMPAC.

AMPAC (Version 2.1)

This molecular orbital package, a product of Dewar's research group, is a new improved version of the original AMPAC, containing the semiempirical Hamiltonians for MINDO/3, MNDO and AM1. It uses the BFGS algorithm for geometry optimization, while for the search of the TS uses the Chain method [9, 40].

The Chain method needs, in order to maintain the connectivity of the path, to restrict each distance between two successive points to a given length which in AMPAC is 0.3 Å.

GAUSSIAN 94

A further development of its previous versions (Gaussian 76, 80, 82, 86, 88–92) [80], Gaussian 94[6] is a connected system of programs for performing *semi-empirical* and *ab-initio* molecular orbital (MO) calculations. Gaussian 88–92 includes:

• *Semi-empirical* calculations using the CNDO, INDO, MINDO/3, MNDO and AM1 model Hamiltonians.

• Automated geometry optimization to either minima or saddle points [15, 20, 22, 44, 74], numerical differentiation to produce force constants and reaction path following [15], and so on.

The option in Gaussian for "Optimizing for a Transition State" is sensitive to the curvature of the surface. In the best case, in which the optimization begins in a region known to have the correct curvature (there is a specific option for this in the menu) and steps into a region of undesirable curvature, the full optimization option (available as a control option for the calculations) can be used. This is quite expensive in computer

6. A version of this program-package, with parts rewritten by Czismadia and coworkers, is also called Monster-Gauss because of the tremendous amount of calculations that can perform as well as being monstrous in length because of its *ab-initio* block.

time but the full Newton-Raphson procedure, already implemented in the program, with good second derivatives at every point will reach a stationary point of correct curvature very reliably if started in the desired region (line searches can be conducted with second derivatives at every point). If a stationary region is not carefully selected, it will simply find the nearest extreme point.

An eigenvalue-following, mode-walking optimization method [74, 81] can be requested by another option (OPT=EF) [43, 82] that is available for both minima and TS, with second, first, or no analytic derivatives as indicated by internal options (CalcAll, CalcFC, default or EnOnly). This choice is often superior to the Berny[7] method, but has a dimensioning limit of variables (50 active variables). By default, the lowest mode is followed. This default is correct when already in a region of correct curvature and when the softest mode is to be followed uphill. Other options of interest, in connection with GOPT and TS calculations, are:

• Freezing Variables During Optimization: Frozen variables are only retained for Berny optimizations.

• Curvature Testing: By default the curvature (number of negative eigenvalues) is checked for the transition state optimization. If the number is not correct (1 for a TS), the job is aborted. Here the search for a minimum will succeed because the steepest descent part of the algorithm will keep the optimization moving downward. On the other hand, a TS optimization has little hope if the curvature at the current point is wrong.

• Murtagh-Sargent Optimization: This method almost always converges slower than the Berny algorithm. It is reliable for minima only.

• Berny or Intrinsic Reaction Coordinate (IRC) method: This is an algorithm designed for finding minima mentioned here because it is often used in the Gaussian

7. Berny stands, with tenderness from the Gaussian people, for Bernard Schlegel.

package. With IRC the reaction path leading down from a TS is examined using the method of Gonzalez and (Berny) Schlegel [15]. In this procedure, the geometry is optimized at each point along the reaction path. All other options which control the details of geometry optimizations can be used with IRC.

Although Gaussian 88–92 has many optimization options that can be used in combination with one another, it will be enough for our purposes to note that this package of programs uses Newton-Raphson, Murtagh-Sargent, Fletcher-Powell and Berny's (by default) methods for optimization, whereas for TS search it uses the Cerjan-Miller algorithm [43] and the Linear Synchronous Transit method (LST) [37].

HONDO

This package written by Michel Dupuis and his co-workers [58] uses algorithms that take advantage of analytic energy derivatives. The Cerjan-Milller algorithm [43] is implemented with an updating of the Hessian matrix. This algorithm has proven efficient provided that a "good" second derivative matrix is used [59] (it has been found that a force constant calculated with a small basis set at the starting geometry is adequate).

As an option, the HONDO program allows the user to use the "Distinguished Reaction Coordinate" approach. This approach consists of a series of optimizations with an appropriately chosen coordinate being frozen at adequately chosen values (for further details see the "Constrained Internal Coordinates" [19, 59] outlined the preceeding chapter).

After this, and an inspection of the potential energy curve, it is possible to guess the TS structure. At this point, a geometry optimization of the guessed TS is suggested by using the BFGS algorithm implemented in the program. The method can be used in conjunction with all SCF wavefunctions, as implemented for the geometry optimization.

All other options on HONDO assume that the TS structure is known as well as the vibrational mode corresponding to the imaginary frequency. Then the gradient norm is minimized [18] to find the nearest extrema.

ACES II (Version 1.0)

This package of programs for performing ab-initio calculations, developed in the early 1990's by Bartlett and co-workers [83], contains geometry optimization algorithms that are all based on the Newton-Raphson method, in which the step direction and size are related to the first and second derivatives of the molecular potential energy. In almost all calculations the exact Hessian is not evaluated but approximated. By default ACES II geometry optimization starts with a very crude estimate of the Hessian in which all force constants for bonded interactions are set to 1 hartree/bohr^2, all bending force constants are set to 0.25 hartree/bohr^2, and all torsional force constants are set to 0.10 hartree/bohr^2. An alternative Hessian is used for some small systems, allowing the use of an . In the search for a minimum, the method implemented in this package can be used when the initial structure is in a region where the second derivative matrix index is nonzero. Moreover, a very efficient minimization scheme, particularly if the Hessian is available, is included in this package of programs, namely, a Morse-adjusted Newton-Raphson search for a minimum.

For the TS search ACES II uses the Cerjan-Miller algorithm [43]. This involves following an eigenvector of the Hessian matrix (that corresponds to a negative eigenvalue) to locate the stationary point, ensuring that it will stay within the region of the TS. Finally, to ensure that a TS has been obtained, the vibrational frequencies are evaluated by taking finite differences.

CHAPTER 4
THE LINE-THEN-PLANE MODEL

Introduction

With the exception of the Synchronous Transit [37] and the Normalization technique [41] models both of which consider the distance between reactants (R) and products (P) while searching in a linear fashion for the TS to finally minimize the maxima found all other procedures discussed in chapter 2 use up-hill methods. Those procedures start from reactants through a second order expansion of the energy in terms of a Taylor series.

For comparison we have collected in Table 2.1 the procedures reviewed in chapter 2, emphasizing their principal features, advantages and, especially, their disadvantages. If we focus on the disadvantages, we notice a general trend in the problems that appear when searching for a TS (which are similar to the geometry optimization ones):

- Costly evaluation of the Hessian matrix.
- Difficulties in identifying an appropriate reaction coordinate.
- Requirement of a good initial guess of the TS.
- Convergence achievement.
- Large number of moves needed (*i.e.* random procedures).

Because of these problems, a better method to find the TS should consider both reactants and products because both contain information on the appropriate saddle point.

Here we present a procedure that is based on a continuous walking from R to P (and vice-versa) with a fixed step length, along a line connecting them. By minimizing the energy at those new points a new line is drawn and the procedure is repeated until a pre-established criterion to find the maxima (or minima) is fulfilled.

The procedure that we present is very simple and has been designed to overcome the problems enumerated above. Thus, we have built up a strategy to find the TS, which makes use of the line search technique, that has the advantage of using a reduced number of calculations, has a simple and convenient expression of projected coordinates, does not require evaluation of the Hessian matrix, considers an intermediate of reaction, and involves the idea of finding the TS(s) starting simultaneously from R and P.

The algorithm does not require the evaluation of the Hessian. As a result, it is much faster in its execution than most of the methods presently in use, and it is applicable for searching the potential energy surfaces of rather large systems. The procedure is not completely unlike the Saddle procedure of Dewar, Healy and Stewart [42] or the line procedure of Halgren and Lipscomb [37], but does differ in a rather substantive way. In this new technique, the line direction is allowed to change during the walk, initially from a line connecting product and reactant to points that represents them. These representative points are determined through minimization of all the coordinates that are perpendicular to the connecting line. The efficiency of this procedure rests upon the observation that it is faster and easier to minimize repeatedly in the N–1 directions than it is to evaluate $N(N + 1)/2$ second derivatives, where N represents the number of variables (coordinates) to be searched.

When the interest is to focus on the shape of the reaction path (mechanism) we suggest, as an alternative simpler strategy, to find the TS(s) so that, when there, the reaction path is constructed by using a down-hill procedure to reactants and products.

However we do not recommend this sequence. We will introduce, unstead, the LTP algorithm and discuss its properties, the step and the speed-up of the procedure by using Hammond's postulate. Finally, Hammond's postulate adapted LTP techniques are discussed.

The *Line-Then-Plane* (LTP) Search Technique

This procedure, originally conceived to find TS in chemical reactions, requires knowledge of reactants (R) and products (P); no previous knowledge of the TS is necessary. It makes partial use of both the line search technique and the search for minima in perpendicular directions that have been discussed already. Figure 4.1 shows the behavior of this technique.

As in the Saddle method of Dewar [40–42] we begin by calculating the structures of the reactant R and product P. A difference vector $\mathbf{d}_i = \mathbf{P}_i - \mathbf{R}_i$ is defined

$$d_i = \left[(X_P - X_R)_i^2 + (Y_P - Y_R)_i^2 + (Z_P - Z_R)_i^2\right]^{1/2} \tag{82}$$

with i = 0, 1, 2, 3, ... and we walk a fraction of the way from $\mathbf{R}_i$ to $\mathbf{P}_i$ along $-\mathbf{d}_i$, and from $\mathbf{P}_i$ to $\mathbf{R}_i$ along $\mathbf{d}_i$. The structures at these two points are minimized in the plane (*i.e.* all directions) perpendicular to $\mathbf{d}_i$, defining the new points $\mathbf{R}_{i+1}$ and $\mathbf{P}_{i+1}$. A new difference vector $\mathbf{d}_{i+1} = \mathbf{P}_{i+1} - \mathbf{R}_{i+1}$ is then defined and the procedure repeated. The steps along $\mathbf{d}_i$ are conservative initially, but increased as a percentage of the norm of $\mathbf{d}_i$, that is $\sqrt{\mathbf{d}_i^\dagger \mathbf{d}_i}$, as the TS is approached. The BFGS technique [18, 21, 26, 27] is used to minimize the energy in the hyperplane perpendicularly to $\mathbf{d}_i$.

At this point we distinguish the Exact and the Approximate LTP procedures:

• Exact LTP: The BFGS technique is used to minimize the energy in the hyperplane perpendicular to the direction $\mathbf{d}_i$.

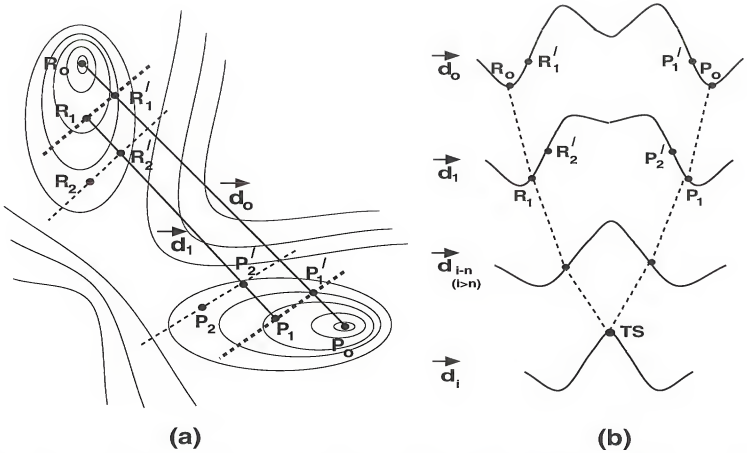


Figure 4.1. (a) General scheme for the Line-Then-Plane (LTP) procedure where the scanning is performed between the last two minimal points found, from both **R** and **P**. (b) A sequential transverse view of the planes containing the projected and the in-line ($\mathbf{d_i}$) points.

• Approximate LTP: No up-date of **H** is performed. For the minimization of the energy, in perpendicular directions to $\mathbf{d}_i$, we use the identity matrix as the Hessian (**PH = P1 = P**) *only once* (steepest descent), so that the projected coordinates depends only on the gradient g_i and on the projector $\mathbf{P}_{d_i}$ (*i.e.*, $i = 0$), see below.

The Algorithm

We adopt the following algorithm, shown graphically in Figures 4.1 a-b:

Step 1) Calculate Reactant and Product geometries $\mathbf{R}_o^p$ and $\mathbf{P}_o^p$ (the super index p stands for projected coordinates). Set counters i = j = 0 (see below).

Step 2) Define the difference vector $\mathbf{d}_i = \mathbf{P}_i^p - \mathbf{R}_i^p$. If the norm $\mathbf{d}_i^\dagger \mathbf{d}_i < \mathrm{T}$ (a given threshold), stop. Otherwise,

Step 3) Examine $\sigma^\dagger(\mathbf{R}_{i-1}^p)\mathbf{d}_i$ and $-\sigma^\dagger(\mathbf{P}_{i-1}^p)\mathbf{d}_i$, where $\sigma(\mathbf{x}) = (\partial E/\partial \mathbf{d}_i)_{\mathbf{x}}$

$$Set \qquad \mathbf{R}_i^l = \mathbf{R}_i^p \quad ; \quad \mathbf{P}_i^l = \mathbf{P}_i^p \tag{83}$$

(the super index l stands for variables *in the line* that connects the corresponding R's and P's) unless,

$$\mathbf{i}) \quad If \quad \sigma^\dagger(\mathbf{R}_i^p)\,\mathbf{d}_i < \mathbf{0} \quad \text{set} \quad \mathbf{P}_i^l = \mathbf{R}_i^p \quad \text{and} \quad \mathbf{R}_i^l = \mathbf{R}_{i-1}^l \tag{84}$$

or,

$$\mathbf{ii}) \quad If \quad -\sigma^\dagger(\mathbf{P}_i^p)\,\mathbf{d}_i < \mathbf{0} \quad \text{set} \quad \mathbf{R}_i^l = \mathbf{P}_i^p \quad \text{and} \quad \mathbf{P}_i^l = \mathbf{P}_{i-1}^l \tag{85}$$

Step 4) Walk along $\mathbf{d}_i$ from $\mathbf{R}_i^l$ to $\mathbf{P}_i^l$, and *vice-versa*, a fixed step of length: $\mathbf{s}_i = \mathbf{d}_i/N_i$

$$Set \quad \mathbf{R}_{i+1}^l = \mathbf{R}_i^l + \mathbf{s}_i \quad and \quad \mathbf{P}_{i+1}^l = \mathbf{P}_i^l - \mathbf{s}_i \tag{86}$$

unless,

$$\begin{aligned} &If \quad T < \mathbf{d}_i^\dagger \mathbf{d}_i \leq \mathbf{s}_j \mathbf{s}_j^\dagger \quad set \quad N_i = N_i/\mathbf{2} \quad and \quad j = i \\ &If \quad N_i < \mathbf{2} \quad set \quad N_i = \mathbf{2} \end{aligned} \tag{87}$$

Step 5) As in Step 3), examine $\sigma^{\dagger}(\mathbf{R}^{l}_{i+1})\mathbf{s}_i$ and $-\sigma^{\dagger}(\mathbf{P}^{l}_{i+1})\mathbf{s}_i$:

$$If \quad \sigma^{\dagger}(\mathbf{R}^{l}_{i+1})\ \mathbf{s}_i < \mathbf{0} \quad \text{set} \quad \mathbf{P}^{l}_{i+1} = \mathbf{R}^{l}_{i+1} \quad \text{and} \quad \mathbf{R}^{l}_{i+1} = \mathbf{R}^{l}_{i} \tag{88}$$

$$If \quad -\sigma^{\dagger}(\mathbf{P}^{l}_{i+1})\ \mathbf{s}_i < \mathbf{0} \quad \text{set} \quad \mathbf{R}^{l}_{i+1} = \mathbf{P}^{l}_{i+1} \quad \text{and} \quad \mathbf{P}^{l}_{i+1} = \mathbf{P}^{l}_{i} \tag{89}$$

Step 6) Minimize in the hyperplanes perpendicular to $\mathbf{d}_i$ containing $\mathbf{R}^{l}_{i+1}$ and $\mathbf{P}^{l}_{i+1}$, to obtain projected points $\mathbf{R}^{p}_{i+1}$ and $\mathbf{P}^{p}_{i+1}$ respectively. Set $i = i+1$ and go to Step 2).

Although one can demonstrate that this procedure must lead to the Transition State on a continuous potential energy surface $E(\mathbf{x})$ if the steps are conservative enough, the norm of the gradient ($\mathbf{g}$) and the Hessian ($\mathbf{H}$) are examined at convergence in order to insure that the converged point on $E(\mathbf{x})$ has the right inertia, (*i.e.* $\mathbf{g} = 0$) and $\mathbf{H}$ has one and only one negative eigenvalue.

To accelerate convergence to the TS, we might add to Step 3) a further test that becomes useful as the TS is approached,

Step 3) iii)

$$If \quad \Delta\sigma^{\dagger}\mathbf{s}(k) = \sigma^{\dagger}(\mathbf{P}^{i}_{k+1})\ \mathbf{s}_k - \sigma^{\dagger}(\mathbf{R}^{i}_{k+1})\ \mathbf{s}_i < T' \qquad and,$$

$$\Delta E(k) = E(\mathbf{P}_k) - E(\mathbf{R}_k) < T'', \quad k = i,\ i-1 \ \ and \ \ i-2 \tag{90}$$

$$Then, \ \ Set: \quad \mathbf{q}_i = (\mathbf{P}_i - \mathbf{R}_i)/\mathbf{2}, \ \ evaluate \ \ E(\mathbf{q}_i),\ g(\mathbf{q}_i) \ \ and \ \ \mathbf{H}(\mathbf{q}_i)$$

$$Else, \ \ If \quad g^{\dagger}(\mathbf{q}_i)g(\mathbf{q}_i) < g^{\dagger}(\mathbf{P}_i)g(\mathbf{P}_i) \quad and \quad -g^{\dagger}(\mathbf{q}_i)\mathbf{d}_i < \mathbf{0} \tag{91}$$

$$Then \ \ q_i \ \ replaces \ \ \mathbf{P}_i \ \ or,$$

$$If \quad g^{\dagger}(\mathbf{q}_i)g(\mathbf{q}_i) < g^{\dagger}(\mathbf{R}_i)g(\mathbf{R}_i) \ \ and \ \ -g^{\dagger}(\mathbf{q}_i)\mathbf{d}_i < 0 \tag{92}$$

$$Then \quad q_i \ \ replaces \ \ \mathbf{R}_i.$$

Then, go to Step 2), *Else*, go to Step 4).

Here $\mathbf{q}_i$ refers to the coordinates that represents a conformation that is very close to the TS structure.

Based on our experience, the choice $N_o = 10$ generates a conservative initial step and suitable thresholds are $T = T' = T'' = T''' = 10^{-4}$ arbitrary units. We have studied the variation of the step size with the number of energy evaluations needed to converge to the saddle point, as is described in a subsequent section.

The strategy delineated above is also successful even for systems which have intermediate structures between **R** and **P**. The tests indicated in equations (84–85) under Step 3) and (88) and (89) under Step 5) disclose potential turning points, caused either by a too large a step from **R** toward **P** or **P** toward **R**.

The reaction path can be approximated by connecting all points $\mathbf{R}_i$ and $\mathbf{P}_i$. An approximate and faster procedure would be to quit in Steps 3) or 5) thereby avoiding the reset of coordinates between consecutive steps.

Then the displacement $\mathbf{d}_i$ can now be divided in smaller parts (say 4) and the procedure continued as before. The last half is now submitted to a perpendicular minima line search founding a last point $\boldsymbol{X}_e$, the TS.

In general, the TS is said to be found if the gradient norm is zero and if the Hessian has one and only one negative eigenvalue, respectively. As for LTP, the transition will be considered to be found when the norm of the displacement vector $\mathbf{d}i$ is smaller than a pre-established convergency threshold T_c (usually $T_c < 10^{-3}$). Nevertheless, the general conditions are checked at the estimated saddle point (*i.e.* $\sigma(\boldsymbol{X}_e) = 0$).

Minimizing in Perpendicular Directions: Search for Minima

The coordinates perpendicular to the direction $\mathbf{d}_i$ are obtained by projection, and the energy in the hyperplane minimized using the BFGS algorithm as developed by Head and Zerner [27]. It has to be pointed out that translations and rotations must be eliminated from $\mathbf{G} = \mathbf{H}^{-1}$ as they represent zero eigenvalues of **H**, in order to construct

a projector free of them. This requirement has been included in the ZINDO program package [26] as part of the implementation of the LTP techniques.

This procedure is restricted to the projected coordinates

$$\mathbf{P}_{\mathbf{d}_i}(\mathbf{x}_{i+1} - \mathbf{s}_{i+1}) = \mathbf{P}_{\mathbf{d}_i}(\mathbf{x}_{i+1} - \mathbf{x}_{i+1}^{(l)}) = -\alpha \mathbf{P}_{\mathbf{d}_i} \mathbf{G}_{i+1} \mathbf{P}_{\mathbf{d}_i} g_{i+1} \tag{93}$$

where

$$\mathbf{s}_{i+1} = \frac{1}{\mathbf{N}} \mathbf{d}_i = \mathbf{x}_{i+1}^{l} \tag{94}$$

is the step (coordinates) along the line connecting projected R's and P's.

In a more compact way, we can write Eqn. (93) as

$$(\mathbf{x}_{i+1} - \mathbf{x}_{i+1}^{(l)})^{\mathbf{P}_{\mathbf{d}_i}} = -\alpha \, \mathbf{G}_{i+1}^{\mathbf{P}_{\mathbf{d}_i}} \, g_{i+1}^{\mathbf{P}_{\mathbf{d}_i}} \tag{95}$$

or

$$\mathbf{x}_{i+1}^{(p_i)} = \mathbf{x}_{i+1}^{(l)} - \alpha \, g_{i+1}^{p_i} \, \mathbf{G}_{i+1}^{p_i} \tag{96}$$

where the projector perpendicular to $\mathbf{d}_i$ is defined as [17]:

$$\mathrm{P}_{d_i} = \mathrm{I} - \frac{\mathbf{d}_i \mathbf{d}_i^{\dagger}}{\mathbf{d}_i^{\dagger} \mathbf{d}_i} \, . \tag{97}$$

In these equations, α is the line search parameter which determines how far along the direction $\mathbf{s_{i+1}}$ of equation (94) one should proceed. For the simple test cases studied in the next chapter, there is but one perpendicular direction, and we set $\alpha = 0.3$ for all i which is a more conservative value than that recommended by Zerner and his collaborators ($\alpha_o = 0.4$, and all other $\alpha_i = 1.0$) [17, 19, 27].

It can be demonstrated easily that, computationally, it is much more convenient to project out only the forces rather than project the forces and the second derivative matrix at the same time. Consequently the new projected coordinates are now obtained as

$$\mathbf{x}_{i+1}^{(p_i)} = \mathbf{x}_{i+1}^{(l)} - \alpha \, \mathbf{G}_{i+1} \, g_{i+1}^{p_i} \tag{98}$$

where the inverse Hessian $\mathbf{G}$ is updated using any appropriate technique.

Projector Properties

The projector $\mathbf{P}_{\mathbf{d}_i}$ must be well behaved (*i.e.* it has to fulfil the conditions of being idempotent and hermitian).

Idempotency: $\mathbf{P} = \mathbf{P}^2$

We start from the definition of the projector:

$$\mathbf{P} = \mathbf{I} - \frac{dd^\dagger}{d^\dagger d} \tag{99}$$

consequently

$$\begin{aligned}
\mathbf{P}^2 &= \left[\mathbf{I} - \frac{dd^\dagger}{d^\dagger d}\right]\left[\mathbf{I} - \frac{dd^\dagger}{d^\dagger d}\right] \\
&= \mathbf{I} - \frac{dd^\dagger}{d^\dagger d} - \frac{dd^\dagger}{d^\dagger d} + \frac{dd^\dagger dd^\dagger}{d^\dagger dd^\dagger d} \\
&= \mathbf{I} - \frac{2dd^\dagger}{d^\dagger d} + \frac{dd^\dagger dd^\dagger}{d^\dagger dd^\dagger d} \qquad (100) \\
&= \mathbf{I} - \frac{2dd^\dagger}{d^\dagger d} + \frac{dd^\dagger}{d^\dagger d} \\
&= \mathbf{I} - \frac{dd^\dagger}{d^\dagger d} = \mathbf{P}
\end{aligned}$$

q.e.d.

Hermiticity: $\mathbf{P} = \mathbf{P}^\dagger$

$$\mathbf{P} = \mathbf{I} - \frac{dd^\dagger}{d^\dagger d} \tag{101}$$

$$\begin{aligned}
\mathbf{P}^\dagger &= \left[\mathbf{I} - \frac{dd^\dagger}{d^\dagger d}\right]^\dagger \\
&= \mathbf{I} - \left[\frac{dd^\dagger}{d^\dagger d}\right]^\dagger = \mathbf{I} - \frac{dd^\dagger}{d^\dagger d} = \mathbf{P}
\end{aligned} \tag{102}$$

q.e.d.

LTP Convergency

Consider the coordinates difference

$$\mathbf{q} = \mathbf{X}_{i+1} - \mathbf{X}^* \quad (< \varepsilon) \tag{103}$$

where $\mathbf{X}^*$ represents a maximum, the TS ($\mathbf{X}^* \equiv \mathbf{X}_{TS}$).

Defining, for the neighborhood of the TS $\mathbf{X}^*$:

$$\lim_{\varepsilon \to 0} \mathbf{q} = 0 \Rightarrow \Theta(\mathbf{q}) = a \text{ iff } \exists\, \mathbf{c} \in \Re / |a| \leq \mathbf{cq} \,. \tag{104}$$

The gradient around a given point $\mathbf{X}_\kappa$:

$$g(\mathbf{X}_\kappa + \mathbf{q}_\kappa) = g(\mathbf{X}_\kappa) + \mathbf{H}_\kappa \mathbf{q}_\kappa + \Theta(|\mathbf{q}_\kappa|^2) \tag{105}$$

but $\mathbf{q}_\kappa = -\mathbf{q}_\kappa$, then:

$$g(\mathbf{X}_\kappa - \mathbf{X}_\kappa + \mathbf{X}^*) = g(\mathbf{X}_\kappa) + \mathbf{H}_\kappa \mathbf{q}_\kappa + \Theta(|\mathbf{q}_\kappa|^2)$$

$$g(\mathbf{X}^*) = g_\kappa - \mathbf{H}_\kappa \mathbf{q}_\kappa + \Theta(|\mathbf{q}_\kappa|^2) \,. \tag{106}$$

If $\mathbf{X}_\kappa$ is too close to $\mathbf{X}^*$ (with $\mathbf{H_k}$ with only one negative eigenvalue), the considered region of the space exists by continuity of the Hessian $\mathbf{H}$. Consequently, the κ th iteration exists.

Projecting from the left by $\mathbf{H}_\kappa^{-1}$:

$$\mathbf{H}_\kappa^{-1} g(\mathbf{X}^*) = 0 = \mathbf{H}_\kappa^{-1} g_\kappa - \mathbf{H}_\kappa^{-1} \mathbf{H}_\kappa \mathbf{q}_\kappa + \Theta(|\mathbf{q}_\kappa|^2) \tag{107}$$

but: $\mathbf{g}_\kappa \mathbf{H}_\kappa^{-1} = -\mathbf{s}_\kappa$, which is a Newton-Raphson like step.

Finally:

$$0 = -s_\kappa - q_\kappa + \Theta(|q_\kappa|^2) = - q'_{\kappa+1} + \Theta(|q_\kappa|^2) \qquad (108)$$

but according to our original definition:

$$\exists\, c \in \Re / |q'_{\kappa+1}| \leq c|q'_\kappa|^2 \ . \qquad (109)$$

If X_κ is very close to X^*, for which:

$$|q| \leq \alpha/c \quad , \quad 0 < \alpha < 1 \qquad (110)$$

by induction, and because $X_\kappa \rightarrow X^*$, the iteration is defined and it exists for all κ and $|q_\kappa| \rightarrow 0$. Consequently, by construction, LTP always will go uphill in the search for a maximum.

To ensure that the new projected points R and P are perpendicular to the reaction coordinate di, we must show now that the energy is a minimum in these directions.

Consider the second-order expansion for the Energy

$$E(x) = E(x_o) + q^\dagger g + \frac{1}{2} q^\dagger H q \ . \qquad (111)$$

From the gradient expansion

$$g = g_o + Hq \qquad and \qquad q = -H^{-1}g \qquad or \qquad g = -Hq \ . \qquad (112)$$

Introducing g in the equation for the energy, we get

$$E(x) = E(x_o) - \frac{1}{2} q^\dagger H q \qquad (113)$$

which demonstrates that the energy in perpendicular directions to the step is minimized by a steepest-descent-like term in which H is positive definite.

From these, we conclude that a second (or higher) order LTP iteration converges.

The Step

A good step will provide a good starting point for the next step, such that the maximization will converge without problems in a reduced number of iterations along the chosen direction. In general, almost all algorithms take their steps without considering previous information about the PES.

In developing LTP, three ways of stepping were studied. The first stepping method is a superimposed step given by a fraction (1/N) of the displacement vector between projected products and reactants coordinates. The second stepping method is based on a proportionality relationship between the actual and previous step. It is shown that this choice will locate the TS (not its final position) at most at half of the size of N, that is, around N/2 LTP cycles, because the final displacements are very small. The stepping method is based on the knowledge of information about the PES given by the current projected point (reactants or products) where the value of N is then estimated by relating the LTP step to the Newton-Raphson one (since LTP is a Newton-Rapson like algorithm).

Default Step

In LTP, the step (s_i) is a fraction of the current displacement vector ($\mathbf{d}_i$)

$$\mathbf{s}_i = \mathbf{d}_i / N \tag{114}$$

where N is a number greater than one. For the first iteration N = 10 (an arbitrary choice suggested after many test calculations) and thereafter the distance between current reactants and products is checked to be not less than a given threshold (say 10^{-3}), otherwise N is reset to 2.5.

Updated Step

A convenient decision on how the step should be taken comes from an algorithm that will decide automatically what the value of N should be for the new LTP cycle once the displacement vector is known. To accomplish this the next step is redefined as to be directly proportional to the previous one:

$$\mathbf{s}_{i+1} = \mathbf{d}_{i+1}/N_{i+1} \propto \mathbf{d}_i/N_i = \mathbf{s}_i \ . \tag{115}$$

Now the problem at hand is an estimation of the value of N_{i+1} and consequently the next move. For this, we consider the following relation between the next $(i+1)$ and the previous (i) steps

$$\mathbf{d}_{i+1}/N_{i+1} = \lambda \mathbf{d}_i/N_i \tag{116}$$

where $0 < \lambda \leq 1$. Projecting now from the left by $\mathbf{d}_i^\dagger$ we obtain,

$$N_{i+1} = N_i \frac{d_i^\dagger d_{i+1}}{d_i^\dagger d_i} \tag{117}$$

when $\lambda = 1$.

Alternately, it might be better to consider a relation with a penalty function on it. This can be written easily as

$$\frac{|d_{i+1}|}{N_{i+1}} = \frac{\lambda |d_i|}{N_i} \left[\frac{(s_i^\dagger s_i)^2}{(\chi_{i_R}^\dagger \chi_{i_P})^2 + (s_i^\dagger s_i)^2} \right] \tag{118}$$

where χ_{i_R} and χ_{i_P} are the difference vectors between the new projected reactant and product, and their corresponding coordinates in the line (step from where the searches start).

Newton-Raphson-Like Step

Consider now the usual LTP step. We want to take a non-arbitrary step based on previous knowledge of the curvature of the region in which we are walking.

Furthermore, we want, at any cycle, the LTP step (s_{LTP}) to be as well behaved as the Newton-Raphson (s_{NR}) one

$$\frac{1}{N}\mathbf{d} = \mathbf{s}_{LTP} = \mathbf{s}_{NR} = \alpha \mathbf{H}^{-1} g \tag{119}$$

where α is a term that comes from the line search technique and the gradient ($\mathbf{g}$) and the displacement ($\mathbf{d}$) are column vectors. Note that only the absolute value of N should be considered. This is because the direction of the walk as defined by the LTP algorithm, is positive when going from reactants to products and is negative in the opposite direction.

Projecting from the left by the gradient complex conjugate ($g^{\dagger}$)

$$\frac{1}{N} g^{\dagger}\mathbf{d} = \alpha g^{\dagger}\mathbf{H}^{-1} g \tag{120}$$

we derive,

$$N = \left(\frac{1}{\alpha}\right) \frac{g^{\dagger}\mathbf{d}}{g^{\dagger}\mathbf{H}^{-1} g} \ . \tag{121}$$

It has been suggested, and shown, by Zerner and his co-workers [17, 19, 27], that for the initial Newton-Raphson step a good choice is to set α = 0.4 and the inverse of the Hessian as the identity matrix. Consequently we can have an approximation to the estimation of N as

$$N = \left(\frac{5}{2}\right) \frac{g^{\dagger}\mathbf{d}}{g^{\dagger} g} \ . \tag{122}$$

This stepping might not be convenient when searching in the vicinity of the saddle point because the denominator will be too small and N will be too large.

Hammond's-Postulate-Adapted LTP Methods

Introduction

It might be argued that LTP, because of its twofold search (reactants and products at the same time), requires too many steps or that it needs twice the amount of effort

(steps) required by other algorithms such as augmented Hessian [19, 48, 49]. Hence, the Augmented Hessian method will be extensively used for comparison. This concern, and the desire to have an algorithm that will move faster and efficiently towards the TS, brought us to the approximate LTP procedure ennunciated in the previous section. However, and by construction, this lack of specific information about the curvature of the potential energy surface provided by the Hessian can be a drawback.

With these problems and goals in mind, we recall Hammond's postulate (HP) [6], which states that the TS will resemble more the initial reactants (R_o) or products (P_o) according to whether the initial or the final state, is higher in energy. However, we have already mentioned some not uncommon examples for which HP fails.

In this section we study the inclusion of HP in order to save some computational efforts by reducing the number of steps. We will do this by adapting LTP to Hammond's postulate and consequently generate two more LTP like procedures, the Hammond-Adapted-Line-Then-Plane procedures (HALTP) and the Restricted HALTP (RHALTP) procedures. For these, the energy of both initial reactants (R_o) and products (P_o) (E_{R_o} and E_{P_o}, respectively) will be considered.

Hammond-Adapted LTP Procedure (HALTP)

Two situations need to be considered:

a) If $\mathbf{E_{R_o} \geq E_{P_o}}$: This is the original (exact and approximate) LTP as described above.

b) If $\mathbf{E_{R_o} < E_{P_o}}$: Reset to a new set of coordinates (prime): $R_o^{'} = P_o$ and $P_o^{'} = R_o$.

The situation described in b) is shown in Figure 4.2, after which LTP will continue as before. This particular situation can also be seen as if the search starts from the

original products. The advantage of this adaptation lies precisely in a reduction on the amount of energy evaluations (LTP cycles) as now LTP will start searching from the geometry of highest energy.

Restricted Hammond Adapted LTP (RHALTP)

In this case the same two situations depicted before are analyzed where the concept of Hammond's postulate is now strictly enforced. The first subcase still leave us with the classical LTP ($E_{R_o} \geq E_{P_o}$), but the second subcase (with $E_{R_o} < E_{P_o}$ as condition) is now modified as follows:

RHALTP I. If ($E_{R_o} < E_{P_o}$) then, do not move the initial products.

This means that the coordinates of the starting products, characterized by P_o, are held constant.

This choice will allow the reactants to move uphill faster towards the TS by being lifted by the products, as shown in Figure 4.3. This possibility is of particular interest when one is concerned with following the path of the reaction under study. The idea is tested in the next chapter for the inversion of ammonia reaction.

RHALTP II. If ($E_{R_o} < E_{P_o}$) then, do not move the initial reactants.

This time we consider that the reactants, characterized by R_o, remain as the initial ones lifting the products towards the TS, as shown in Figure 4.4. The idea is tested, again in chapter 6, for the non-symmetric inversion of ammonia reaction.

It has to be pointed out that RHALTP I and RHALTP II are not the same procedure with different label for reactants and products (and of technique), because the energetics of the changed coordinates are completely different.

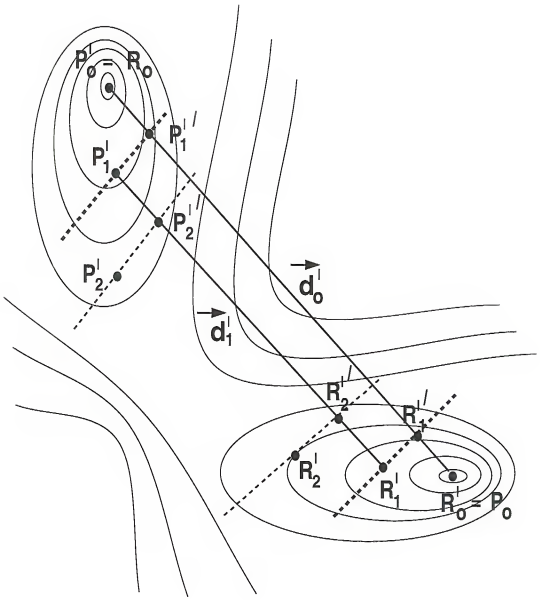


Figure 4.2. Hammond-Adapted-Line-Then-Plane (HALTP) technique in which the search starts from the set of coordinates of higher energy according to: $E_{R_o} < E_{P_o}$.

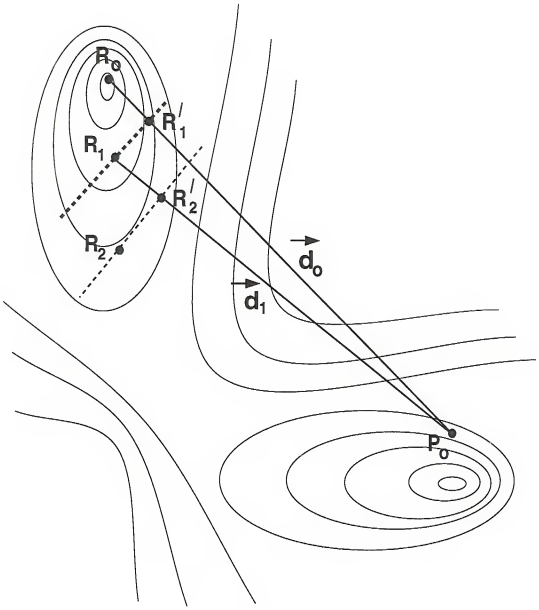


Figure 4.3. Products Restricted-Hammond-Adapted-Line-Then-Plane technique (RHALTP I). The coordinates of products characterized by P_o are held constant, lifting the reactants towards the TS.

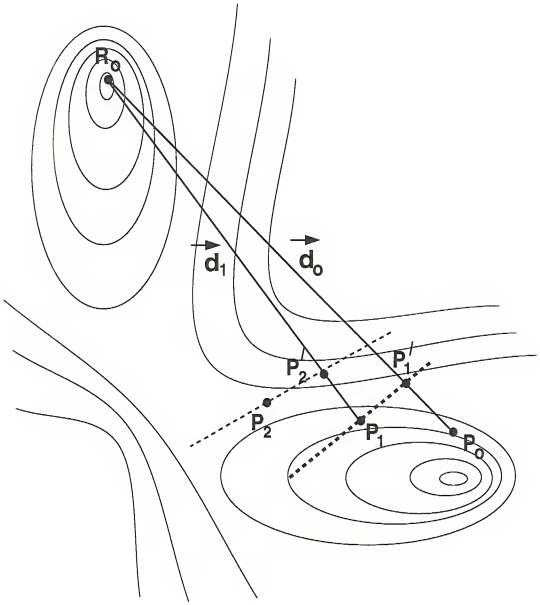


Figure 4.4. Reactants Restricted-Hammond-Adapted-Line-Then-Plane technique (RHALTP II). The coordinates of reactants characterized by P_o are held constant, lifting the products towards the TS.

CHAPTER 5
GEOMETRY OPTIMIZATION

Introduction

Almost all the procedures discussed in chapter 2 use Steepest Descent methods to search for minima through a second-order expansion of the energy in terms of a Taylor series.

From the Geometry Optimization procedures reviewed in chapter 2, the general behavior of problems in the search of minima becomes clear:

- Costly evaluation of the Hessian matrix.
- Large number of moves needed (*i.e.* random procedures).
- Convergence problems.

Although several procedures are available, there are still other problems, such as the loss of information about the curvature when the Hessian is not considered. In this way, and as is the case for TS search, the development of new techniques will rely on experimentation, namely that the model must show acceptable behavior on a variety of test functions, chosen to represent the different features of a typical problem.

Because of these problems, it seems that a better method to find the minima (hopefully the global minimum) must consider the initial geometry plus a generated second

one (only at the initial step). Therefore, in addition to position and/or displacement vectors, the displacement vector between the two initial points should also be considered.

A procedure that is based on the Line-Then-Plane technique (LTP), that is, a continuous walking from the lowest energy point through a line connecting the two lowest energy points, is proposed. By minimizing the energy at the new point a new line is drawn between the new point and the one from which the projection was performed. The procedure is repeated until a preestablished criterion to find the minimum is fulfilled. Figure 5.1 illustrates the behavior of this procedure.

The same features already described for TS search with LTP are valid here, that is, this is a procedure that does not require the evaluation of the Hessian. As a result, the proposed method is much swifter in its execution than most of the methods used today, and is applicable for searching for minima in potential energy surfaces of rather large systems. In this technique, the line direction is allowed to change during the down-hill walk, initially from a line connecting the starting geometries that *represent* them. These points are determined through minimization of all the coordinates that are perpendicular to the connecting line. The efficiency of this procedure rests upon the observation (as for LTP), that it is quicker and easier to minimize repeatedly in the N–1 directions than it is to evaluate N(N + 1)/2 second derivatives, where N represents the number of variables (coordinates) to be searched.

ARROBA: A ***Line-Then-Plane*** Geometry Optimization Technique

This procedure requires a single input geometry from which a second set of coordinates will be generated only in the first step. The down-hill walk starts by determining the lowest energy point, making partial use of the line search technique

and the search for minima in perpendicular directions. As introduced previously, a new projected minima is then found. Figure 5.1 illustrates the behavior of this idea.

As in the "Amoeba," or "Simplex" method of Nelder and Mead [30], we begin by calculating the structures of the initial point and a second one generated as:

$$q_2 = q_1 + \beta \tag{123}$$

where β is a 3N dimensional unitary vector (where N = number of atoms) scaled by three different factors (ϕ, χ and ψ, for the x, y and z components, respectively), one can make β a constant (but $\beta \neq 0$). Any of these choices will be the initial guess for the problem and will depend on the size of the system.

Once a second initial point is generated, energies (E) and gradients ($\mathbf{g}$) are evaluated for both initila points ($\mathbf{q_1}$, E_1, g_1 and $\mathbf{q_2}$, E_2, g_2). The strategy then is as follows:

A difference vector $\mathbf{d}_i = \mathbf{q}_2 - \mathbf{q}_1$ is defined

$$d_n = \left[(\mathrm{X}_{i+1} - \mathrm{X}_i)_n^2 + (Y_{i+1} - Y_i)_n^2 + (Z_{i+1} - Z_i)_n^2 \right]^{1/2} \tag{124}$$

with n = 0, 1, 2, 3, ...

The structure of lowest energy of these two points will be minimized in the plane perpendicular to $\mathbf{d}_i$, defining a new point $\mathbf{q}_{i+2}$ and i is reset to $i = i + 1$. A new difference vector $\mathbf{d}_i = \mathbf{q}_{i+1} - \mathbf{q}_i$ is then defined and the procedure is repeated. The norm of $\mathbf{d}_i$ (*i.e.* $\sqrt{\mathbf{d}_i \mathbf{d}_i^{\dagger}}$), is checked for convergence as the minimum is approached. The **BFGS** technique [18, 21, 26, 27] is used to minimize the energy in the hyperplane perpendicular to $\mathbf{d}_i$.

As for the search for maxima, we differentiate between the Exact and the Approximate ARROBA procedures:

• Exact ARROBA: The BFGS technique is used to minimize the energy in the hyperplane perpendicular to the direction $\mathbf{d}_i$.

• Approximate ARROBA: No up-date of **H** is performed. For the minimization of the energy, in directions perpendicular to $\mathbf{d}_i$, we use the identity matrix as the Hessian (**PH = P1 = P**) only once (constrained steepest descent), so that the projected coordinates depend only on the gradient $\mathbf{g}_i$ and on the projector $\mathbf{P}_{d_i}$ (*i.e.*, $i = 0$), as described below.

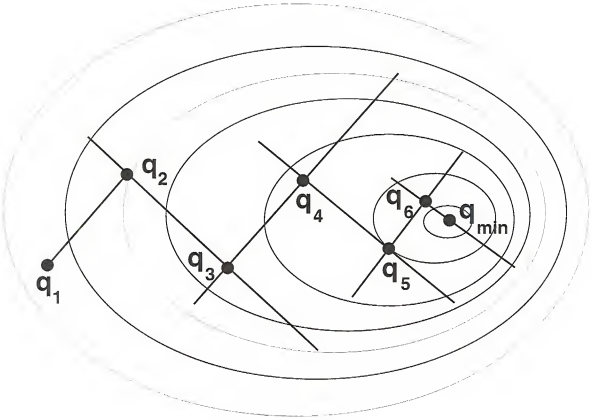


Figure 5.1. Schematic representation of ARROBA, an adapted Line-Then-Plane technique for geometry optimization. The input coordinates ($\mathbf{q_1}$), the initially generated one ($\mathbf{q_2}$), the general zigzag behavior of the procedure and the found minima ($\mathbf{q_{min}}$) are shown.

The Algorithm

We adopt the following algorithm, shown graphically in Figure 5.1:

Step 1) Calculate initial and new generated geometries $\mathbf{q_1}$ and $\mathbf{q_2}$. Set counter $i = 1$.

Step 2) Define the difference vector $\mathbf{d}_i$ for which its norm $\mathbf{d}_i^\dagger \mathbf{d}_i$ is greater than T (a given threshold), else the program will stop:

$$\text{i)} \quad If \quad E_{i+1} < E_i \quad \text{then} \quad \mathbf{d}_i = \mathbf{q}_{i+1} - \mathbf{q}_i \tag{125}$$

$$\text{ii)} \quad If \quad E_{i+1} > E_i \quad \text{then} \quad \mathbf{d}_i = \mathbf{q}_i - \mathbf{q}_{i+1} \tag{126}$$

Else: Stop, and check for convergency: $\mathbf{d}_i^\dagger \mathbf{d}_i > \text{T}$.

Step 3) Minimize in the hyperplanes perpendicular to $\mathbf{d}_i$ containing E_{i+1} to obtain projected points E_{i+2}. The point from where the perpendicular minimization starts is that one with the lowest energy. We set $i = i + 1$, accept the new point if and only if $E_{i+2} < E_{i+1}$, Else: go to Step 2).

Step 4) The new projected point coordinates are given by

$$If \quad E_{i+1} < E_i \quad then \quad \mathbf{q}_{i+2}^p = \mathbf{q}_{i+1}^p - \alpha \, g_{i+1}^p \, \mathbf{G}_{i+1}^p \tag{127}$$

where the upper script (p) stands for projected variables using the projector as showed in the preceding chapter.

It can be shown that this procedure must lead to a minimum that, according to its location in the PES, might be a local or a global minimum, provided the surface is continuous. The norm of the gradient and the Hessian are examined at convergence in order to ensure that the converged point on E($\mathbf{q}$) has the right inertia. The minimum is said to be found if the gradient norm is zero and if all the Hessian eigenvalues are positive.

Minima in Perpendicular Directions

As in the LTP method, ARROBA uses coordinates perpendicular to the direction $\mathbf{d}_i$, obtained by projection. The energy in the perpendicular hyperplane is then minimized using the BFGS algorithm as developed by Head and Zerner [27].

Again, it is computationally much more convenient to project only the forces rather than project them and the Hessian matrix at the same time. Consequently the new projected coordinates are obtained as

$$\mathbf{x}_{i+1}^{(p_i)} = \mathbf{x}_{i+1}^{(l)} - \alpha\, g_{i+1}^{p_i}\, \mathbf{G}_{i+1} \tag{128}$$

where the inverse Hessian is now updated using any appropriate technique.

Convergency

To ensure that the new projected points R and P are perpendicular to the reaction coordinate **d**, we must show that the energy is a minimum in these directions.

Again consider the second order expansion for the energy

$$E(\mathbf{x}) = E(\mathbf{x}_o) + \mathbf{q}^{\dagger} g + \frac{1}{2}\, \mathbf{q}^{\dagger}\mathbf{H}\mathbf{q} \tag{129}$$

and the gradient expansion

$$g = g_o + \mathbf{H}\mathbf{q} \tag{130}$$

and

$$\mathbf{q} = -\mathbf{H}^{-1} g \qquad or \qquad g = -\mathbf{H}\mathbf{q} \tag{131}$$

which is the quasi-Newton condition. Introducing $\mathbf{g}$ in the equation for the energy we get

$$E(\mathbf{x}) = E(\mathbf{x}_o) - \frac{1}{2}\, \mathbf{q}^{\dagger}\mathbf{H}\mathbf{q} \tag{132}$$

which demonstrates that the energy in perpendicular directions to the step is minimized by a steepest-descent-like term in which **H** is positive definite.

From these considerations, we conclude that a second (or higher) order for the ARROBA iterations converges. This minimization procedure has the advantage of using a reduced number of calculations, particularly in the case of the Approximate technique, and does not use the Hessian. It is guaranteed to step down-hill. The minimum found will be a local minima. The search for the global minimum is discussed below.

A Proposed Global Minima Search Algorithm

As discussed in chapter 1, when looking for minima it is very desirable for a procedure to be able to find the global minimum, especially for large molecular systems (proteins, enzymes) for which the most widely used current procedure is the Monte Carlo model often requiring thousands of energy evaluations.

Here we propose a procedure that will have a behavior like Monte Carlo, but does not depend on the temperature and that does not need as many calculations. It uses a jump-out technique, as the warm-up part of the Monte Carlo techniques to take the system out of the local well in which it is trapped.

The algorithm requires a control option from the input file that allows the user to perform several ARROBA calculations. The strategy is as follows:

1) Make an ARROBA minima search.

2) Set counter $i = 1$ and label the new minimum as: $\mathbf{q}_i = \chi_n$ (n is the internal ARROBA counter).

3) Construct a displacement vector ($\mathbf{r}_i$) between the minimum found and the input geometry χ_o:

$$\mathbf{r}_i = \mathbf{q}_i - \chi_o \,. \tag{133}$$

4) Get a new displacement vector $\mathbf{r}_i^{'}$ orthogonal to $\mathbf{r}_i$, that is, in braket notation:

$$\langle r_i^{'} | r_i \rangle = 0 . \tag{134}$$

5) Obtain a new initial set of coordinates $\chi_o^{'}$:

$$\chi_o^{'} = \mathbf{r}_i + \chi_o . \tag{135}$$

6) Check for maximum allowed number of searches M:

$$If \quad i < M \quad Go\ To \quad Step\ 1 \quad Else \quad Stop \tag{136}$$

where M is a pre-established maximum number of iterations.

CHAPTER 6
APPLICATIONS

Introduction

The ideas discussed in the previous chapters have been tested by two different approaches. One approach involves the use of two-dimensional model potential functions to test the behavior of the LTP procedures and compare the results with reports on other methods in the literature [14]. Six model potential functions are examined. The Hammond-Adpated LTP technique has also been tested on three of these functions and, the Restricted-Hammond-Adapted procedures were investigated on a 7^{th} potential function. Finally, using the potential functions, the LTP accuracy and convergence dependence on the step size have been studied.

The LTP method was also tested on several molecular systems: the inversion reaction of water, the symmetric and the non-symmetric isomerization reactions of ammonia, a rotated inversion reaction of ammonia, the hydrogen cyanide isomerization reaction, the formic acid 1,3 sigmatropic shift reaction, the methyl imine isomerization and the thermal retro [2+2] cycloaddition reaction of Oxetane. The accuracy has also been examined in terms of the step and number of energy evaluations required to find the TS in the molecular examples.

For the study of those systems, the Intermediate Neglect of Differential Overlap (INDO) technique [84] has been used, at the Restricted-Hartree-Fock (RHF) level [85] within the ZINDO program package [26]. The minimization procedures are, of course,

limited to no particular energy function, provided it is continuous. The results were compared with those of the Augmented Hessian (AH) technique that uses the same INDO Hamiltonian but evaluates the Hessian at each iteration. All the above-mentioned procedures were implemented in the ZINDO [26] suite of programs.

Model Potential Functions for Transition State

LTP procedures have been tested on six model potential functions which are traditionally used to examine TS searching procedures. The first two, the Halgren-Lipscomb and Cerjan-Miller potential functions, have their TS located closer to the reactant than to the product. The next two potential surfaces, the Hoffman-Nord-Ruedenberg and Culot-Dive-Ghuysen, have the TS located closer to the products than the reactants. The fifth potential function has a TS located midway between reactants and products, and the sixth PES has a steep minimum located in the products region. The results of these tests are collected in Tables 6.1, 6.2 and 6.3 and are discussed below.

The Halgren-Lipscomb Potential Function

The Halgren-Lipscomb potential function [37, 39]:

$$E_{\mathrm{HL}}(x,y) = \left[(x-y)^2 - (5/3)^2\right]^2 + 4(xy-4)^2 + x - y \qquad (137)$$

has two minima (we have chosen points (1.328, 3.012) and (3.0, 1.333) for reactants and products respectively) and one first order saddle point (2.0, 2.0). Figure 6.1 shows the shape of this surface as well as the points obtained with the LTP procedure. Notice that both LTP procedures, Exact and Approximate, walk uphill using the same points.

The Cerjan-Miller Potential Function

Cerjan and Miller's function [43]:

$$E_{\mathrm{CM}}(x,y) = (a - by^2)x^2 \exp(-x^2) + \frac{c}{2}y^2 \qquad (138)$$

has two symmetric TSs located at points (± 1, 0) and a minimum at point (0, 0). As R and P coordinates we have selected points (0,0) and (2.7, 0.05) respectively.

For this procedure, an accurate Hessian is required. Others that have used this function include Simons et al. [31], with a Fletcher-based surface algorithm [22], Banerjee *et al.* [32], with a rational function optimization algorithm, and Abashkin and Russo [86], with a constrained optimization procedure. All of these previous studies have used $a = b = c = 1$ with the exception of Simons and his coworkers, who used $a = c = 1$, $b = 1.2$. Figure 6.2 shows the behavior of our procedure when applied to this potential energy surface. Here both procedures walk towards the TS and are very close to each other (notice the scale on the Y axis).

The Hoffman-Nord-Ruedenberg Potential Function

Hoffman et al. [52] have used the model surface function:

$$E_{\mathrm{HNR}}(x, y) = (xy^2 - yx^2 + x^2 + 2y - 3)/2 \qquad (139)$$

to test their gradient extremal procedure. This function has also been tested by Schlegel [10]. As other algorithms, previously mentioned, these methods require an accurate evaluation of the second derivatives.

The function has two saddle points TS_1 (- 0.8720, 0.7105) and TS_2 (3.1352, 1.2487). In order to test the LTP procedures the points (1, 1) and (5.4980, 1.2874) have been chosen as the R and P coordinates respectively. From these points, a walk towards the TS has been performed. Figure 6.3 shows the behavior of our suggested procedure in this potential function surface.

We note the somewhat chaotic behaviour of the Approximate procedure in the products region, due to its inherent lack of information of the quadrature of the surface.

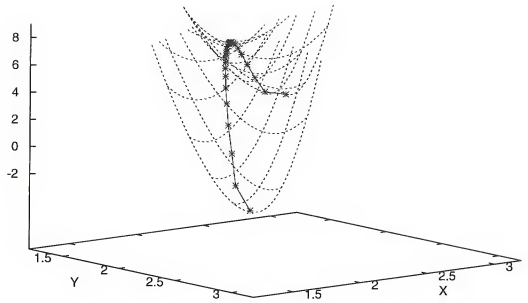


Figure 6.1. The Halgren-Lipscomb Potential Energy Surface $E_{\text{HL}}(x,y) = \left[(x-y)^2 - (5/3)^2\right]^2 + 4(xy-4)^2 + x - y$ showing the full procedure. The Exact and Approximate LTP procedures are represented by crosses and asterisks respectively, walking together (same coordinates) towards TS. The TS is represented by the point where R and P meet.

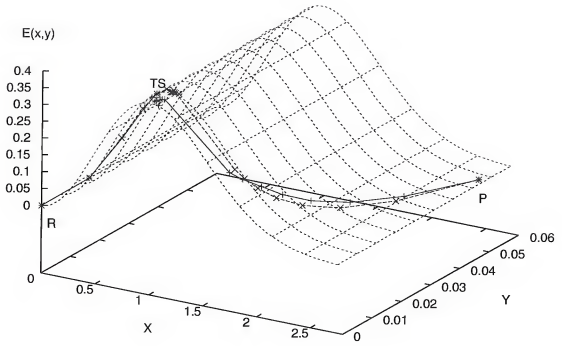


Figure 6.2. Cerjan-Miller $E_{\text{CM}}(x,y) = (a - by^2)x^2 \exp(-x^2) + cy^2/2$ potential energy function, with $a = b = c = 1$. Here the transition state (black diamond for Exact LTP and bold cross for the Approximate LTP) in the 3 dimensional surface is located in the reactants region. The Exact and Approximate LTP methods (represented by crosses and asterisks) satisfy the threshold positions before reaching the same point (see text and Table 6.2).

On the other hand, what seems to be a gap on steps in the uphill walk towards the transition state, is nothing but the convergence acceleration mechanism at work.

The Culot—Dive—Nguyen—Ghuysen Potential Function

Proposed by Culot et al. [11], this test function has, in the range interval [-5, 5], 1 maximum, 4 minima and 3 first-order Saddle Points (points (1); (2), (3), (4), (5); (6), (7) and (8) respectively, identified as in reference [11]). LTP was tested using as R, P and TS points (3), (2) and (7) whose respective coordinates are (3.585, –1.850), (3.0, 2.0) and (3.4, 0.1). The function is:

$$E_{\text{CDNG}}(x,y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2 \qquad (140)$$

In Figure 6.4 the behavior of our procedure for this potential surface is represented. As was the case for the Cerjan-Miller surface, both procedures, Exact and Approximate, walk towards the TS and have in common the same projected points.

A Midpoint Transition State Potential Function

A simple function:

$$E_{\text{MPTS}}(x,y) = \left[(x - 1)(x - 2)\right]^2 + \left[(y - 1)(y - 2)\right]^2 \qquad (141)$$

was used in order to test the ability of our procedure to step up-hill towards the TS in a flat PES. This function has two minima: (1.0, 1.0) used as reactants (R) and (0, 2.0) as products (P) . The TS is located in the midpoint between reactants and products (1.5, 1.5) with a potential energy of 0.125 (arbitrary units). The behavior of our algorithm on this potential surface is shown in Figure 6.5, where it can be seen that both techniques walk towards the TS using the same projected points.

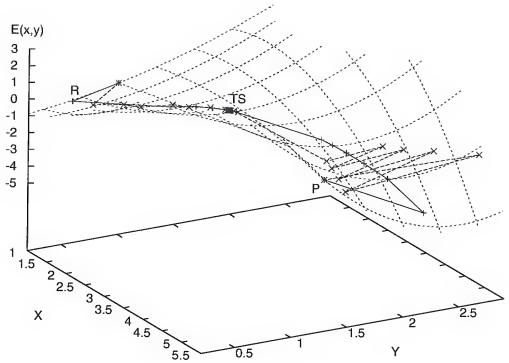


Figure 6.3. The Hoffman-Ross-Ruedenberg potential energy function $E_{\mathrm{HNR}}(x,y) = (xy^2 - yx^2 + x^2 + 2y - 3)/2$. Exact LTP (cross) and Approximate LTP (asterisk) approach the TS (bold square) located in the products region, although the steps of the approximate method deviate considerably from the reaction coordinate.

A Potential Function with a Minimum

This potential energy surface:

$$E_{\text{Min}}(x, y) = x^4 + y^4 + 2xy \tag{142}$$

has the property of having a very steep minimum located at (0.7071 , –0.7071), very close to the starting products coordinates (1.0, –1.0). The reactant is located at (0.2, 0.0). Figure 6.6 demonstrates the downhill behavior of our procedure.

The Approximate procedure starts with erratic behaviour on the products side, which is corrected by the reactants side as the minimum is approached. Here both techniques walk in their own fashion showing their individual characteristics.

Summary of Results

Table 6.1 summarizes the number of function evaluations required by the LTP procedures along with the results of other authors. The comparison is based in the simple "counting" (when possible) of the number of points on the appropriate figures of the corresponding works, assuming that these are the number of steps required to find the TS. Table 6.2 shows our results for the exact LTP as well as the values obtained using the Approximate LTP procedures.

It should be remarked that a better answer is obtained when a fraction of the smallest slope between reactants and products is used as an initial value of α for both procedures.

Because of the nature of the steps advocated, the TS will never be missed, or passed, provided that the steps along the vector connecting $\mathbf{P}_i$ and $\mathbf{R}_i$ are small enough. A clear demonstration of this is given graphically in Figure 4.2b (chapter 4), where it can be seen that the last point, the saddle point: **TS**, is a unique point at which $\mathbf{R}_i = \mathbf{P}_i$. Moreover, it can be demonstrated easily that, because of the Newton-Raphson nature of the step chosen [14], the LTP procedure will always converge to the TS.

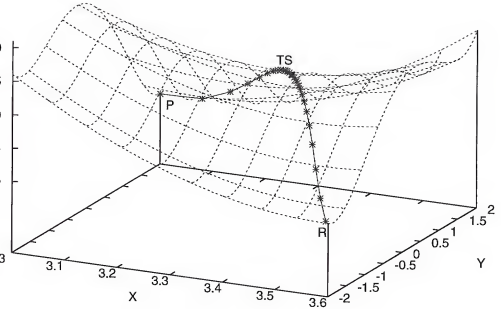


Figure 6.4. Culot's *et. al.* $E_{\text{CDNG}}(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$ potential energy function. Here the is represented by the highest in energy bold asterisk. Notice how both (Exact (crosses) and Approximate (asterisks)) LTP procedures walk uphill with the same coordinates and direction. The TS is located in the Products region.

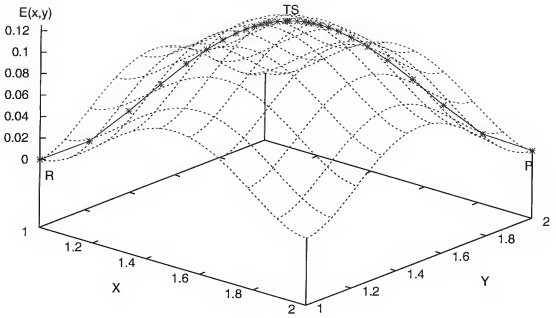


Figure 6.5. Representation of $E_{\text{MPTS}}(x,y) = \left[(x-1)(x-2)\right]^2 + \left[(y-1)(y-2)\right]^2$ potential energy function. The TS is located exactly in the midpoint between reactants and products. Again, both procedures (crosses for the Exact LTP and asterisks for the Approximate LTP) walk uphill together using the same coordinates and direction.

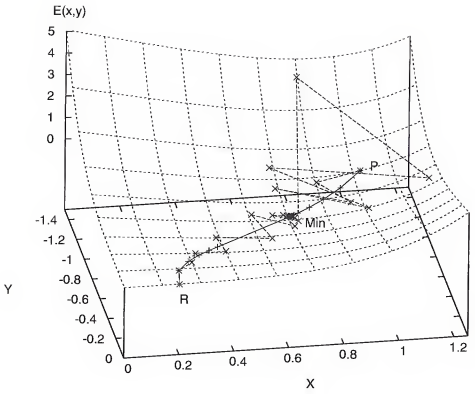


Figure 6.6. Potential energy function with a minima $E_{\text{Min}}(x, y) = x^4 + y^4 + 2xy$. Here the minimum Min (bold bullet) is found in the products region. Again, the continuous line represents the path followed by the Exact LTP procedure whereas the dashed line stands for the Approximate procedure path. Note: the Approximate LTP (X) deviates considerably from the intrinsic reaction coordinate, but does find the TS as does the Exact LTP (crosses), see text.

Table 6.1. Number of function evaluations (NFE) required to find the TS for Cerjan-Miller [43] (CM) and Hoffman-Nord-Ruedenberg [12] (HNR) potential function surfaces (see text). Values in parentheses are for the Approximate LTP procedure.

Model		NFE_{CM}
Lagrange Multipliers [43]		8
Soft-Mode Analytical Hessian [31]		12
Soft-Mode Updated H [31]		23
Stiff-Mode Analytical Hessian [31]		6
Stiff-Mode Updated H [31]		7
C.G. + q-N.M.[a] [87]		10
RFO[b] [25a]		14 and 13
RFO[b] [32]		13[c]
P-RFO[d] [32]		13
RFO + H Updated [32]		16
P-RFO + H Updated [32]		15
Constrained Opt. Tech. [9]		12
Gradient Extremal [10] (HNR)		11
This Work :	NFE_{CM}	20 (22)
	NFE_{HNR}	20 (26)

a. Conjugated Gradients + quasi-Newton Minimization methods.
b. Rational Function Optimization.
c. a = c = 1, b = 1.2.
d. Partitioned RFO.

In practice, we are maximizing along a line $\mathbf{d}_i$, and it is easy to insure that $\partial^2 E/\partial d_i^2$ is **negative** when $\mathbf{R}_t = \mathbf{P}_t$. The perpendicular searches utilize the BFGS procedure starting with a positive Hessian. Since this procedure cannot change the signature of the Hessian it will either minimize in all perpendicular directions (all other $\partial^2 E/\partial x^{(p_i)2} > 0$) or it will fail.

Table 6.2. Coordinates and potential energies for the Exact and Approximate LTP procedures (first and second rows respectively), compared with the Expected results for the TS for the six potential function surfaces (gradient norm less than 10^{-4}, α = 0.3 with N_o = 10). Also displayed are the number of steps used to find the TS and its location (region[c]). The average deviations for these procedures are Exact LTP: Δx = 0.0013, Δy = 0.0037 and ΔE = 0.0018; Approximate LTP: Δx = 0.0015, Δy = 0.0043 and ΔE = 0.0018 (arbitrary units).

	X		Y		E(x,y)	Steps-
	Expect	LTP	Expect	LTP	Expect	LTP	Region[c]
E^a_{HL}	2.0000	1.9955	2.0000	2.0045	7.7161	7.7066	14-R
		1.9955		2.0045		7.7066	14
E_{CM}	1.0000	1.0003	0.0000	0.0024	0.3679	0.3679	10-R
		0.9983		0.0069		0.3679	11
E_{HNR}	3.1352	3.1358	1.2487	1.2489	0.9707	0.9707	10-P
		3.1353		1.2487		0.9707	13
E^a_{CDNG}	3.3852	3.3834	0.0739	0.0879	13.3119	13.3105	15-P
		3.3834		0.0879		13.3105	15
E^b_{MPTS}	1.5000	1.5000	1.5000	1.5000	0.1250	0.1250	15-MP
		1.5000		1.5000		0.1250	15
		1.5000		1.5000		0.1250	4
E_{Min}	0.7071	0.7078	-0.7071	-0.7061	-0.5000	-0.5000	10-P
		0.7079		-0.7070		-0.5000	13

a. Only E_{HL} shows an appreciable error for the threshold, 0.0095 units. For the function value, only E_{CNDG} shows an appreciable percentage error in y of only 0.014.

b. For E_{MPTS} the third row shows results obtained using the simple test according to condition iii on Step 3), whose principal feature is the reduced number of steps (4) used to find the Transition State.

c. R = Reactants, P = Products, MP = Midpoint between initial reactants and products.

On the other hand, it can be seen from Figures 6.1–6.6 that both the exact and the approximate LTP procedures, walk uphill in the same direction with similar

coordinates. In particular, the coordinates are the same for both approaches in Halgren and Lipscomb's, Culot's *et. al.* and MPTS potential functions, almost the same for Cerjan and Miller's function, close for Ruedenberg's *et. al.* functions, and a little bit erratic for the function that presents a minimum very close to the products region (Figure 6.6). In this case the jumpy behavior of the approximate LTP procedure is caused by a Hessian that is far from the projector itself. In spite of this, the results are remarkably good. We may speculate that the Approximate LTP procedure may not accurately follow an intrinsic reaction path initially but will become more accurate as the TS is approached and, because of the size of the step, it will find the TS.

Finally, large separations between consecutive points, as in Cerjan-Miller's (Figure 6.2), Hoffman, Nord and Ruedenberg (Figure 6.3) and the Function with a Minima (Figure 6.6) potential energy functions, are due to the accelerating convergence conditions to the TS, established in Steps 3) and 5) of the algorithm. As for the case of the Hammond-Adapted-Line-Then-Plane technique, we found this, as expected, to result in a small reduction of the number of steps that are needed to find the TS, because it just decides from which initial set of coordinates (reactants or products) the search has to start.

We have tested the Restricted-Hammond-Adapted-LTP procedures on 3 of the preceding potential energy functions, (*i.e.* Cerjan and Miller, Hoffman-Nord-Ruedenberg and Culot-Dive-Nguyen-Ghuysen). The results of these tests are shown in Table 6.3. The reduction in number of steps expected is small, when compared with previous calculations (Table 6.2), because this is a very simple and modest improvement of the algorithm. In fact, the changes are only noticed at the beginning of the search and at each time that an LTP acceleration in the up-hill direction is performed (changes in slopes for example).

Table 6.3. Coordinates and potential energies for both Restricted-Hammond-Adapted-Line-Then-Plane (RHALTP) procedures (Exact and Approximate, models, first and second rows for each potential function, respectively), compared with the Expected results for the TS for three selected potential function surfaces (gradient norm less than 10^{-4}, α = 0.3 with N_o = 10). Also displayed are the number of steps used to find the TS. The average deviation for the Exact LTP procedure for this threshold is Δx = 0.0014, Δy = 0.0044 and ΔE = 0.0001 (arbitrary units), whereas for the Approximate LTP procedure is Δx = 0.0023, Δy = 0.0068 and ΔE = 0.0000 (arbitrary units).

	X		Y		E(x,y)		Steps
	Expect	LTP	Expect	LTP	Expect	LTP	
E^a_{CM}	1.0000	0.9994	0.0000	0.0124	0.3679	0.3679	9
		0.9994		0.0172		0.3679	9
E_{HNR}	3.1352	3.1369	1.2487	1.2490	0.9707	0.9707	18
		3.1288		1.2456		0.9707	14
E_{CDNG}	3.3852	3.3870	0.0739	0.0734	13.3119	13.3121	16
		3.3852		0.0739		13.3119	16

a. For the independent coordinates, for the function value only E_{CM} shows an appreciable percentage error, though the error in y is only 0.0124 and 0.0172 for the Exact and Approximate LTP procedures, respectively.

The Step

As discussed in the preceding chapter, we have studied different ways of stepping in order to get better insight of the potential energy surface and, in this way be capable of choosing the best response from our techniques, as for all of them the stepping will be the same.

We have applied these ideas to the Midpoint Symmetric Potential Function:

$$E_{\text{MPTS}}(x,y) = \left[(x-1)(x-2)\right]^2 + \left[(y-1)(y-2)\right]^2 \qquad (143)$$

Table 6.4. Estimation of the size of the step based on information of the previous one as established on Chapter 6, Case 2 (see text). The results come from applying this stepping to the midpoint symmetric potential function ($E_{\mathrm{MPTS}}(x,y)$). Notice that the size of the step (<s_i>) has been kept almost constant no matter the size of the displacement.

i	N_i	<d_i>	<s_i>	<$d_i d_i$>	<$d_i d_{i+1}$>
1	10.00	14.04	1.72	197.00	
2	8.10	11.39	1.72	129.65	159.6
3	6.04	8.51	1.40	72.45	96.68
4	3.98	5.61	1.72	31.52	47.78
5	1.98	2.84	1.44	8.05	15.68

and summarize the results of the application of both relations derived from Case 2 of the preceding chapter in Table 6.4. It is interesting to notice that the sizes of the steps are approximately the same for each iteration. But what is much more interesting is that, according to the results the TS should be found at the 5th iteration, because the last iteration established that $N_{i+1} \simeq 2$. For any size of the initial N, the number of LTP cycles required to find the TS will always be N/2, provided that N is within a safe range of convergence to the saddle point, and that the potential energy surface has a quadratic behavior.

The same test was performed on Cerjan Miller's potential energy function:

$$E_{\mathrm{CM}}(x,y) = (a - by^2)x^2 \exp(-x^2) + \frac{c}{2}y^2 \qquad (144)$$

for which the step was kept constant. The results are displayed in Table 6.5. We notice that at the beginning of the 6th LTP cycle (until the 11th cycle) the step size is set to 2.5, after the updating of it (N_5) gets closer to 2. At the 6th cycle, an until the 11th one, the coordinates are relabeled because of the acceleration to convergency

properties of LTP, narrowing the region to be examined until at the 12th cycle reactants step the TS is found. The behavior of this alternative stepping is compared against the classical LTP (Exact) by plotting the trajectories followed by both stepping cases, as shown in Figure 6.7.

From these results it becomes clear that there is no saving in steps and that both techniques starts walking using the same points until the 3rd cycle, after which they slightly separate, suffer from the same acceleration techniques, and find the same TS at which they meet. Consequently, it can be assumed that a step factor of 1/10 is as good as any other one, and it will used as default from now on.

Step Size Dependence

The dependence of the convergence of LTP to the TS as a function of the step size has been studied also. For each case the requirements to locate the TS are the same, that is, size of the displacement less than 10^{-3}, maximum component of the gradient less than 10^{-3} and continuous check of the Hessian signature until one negative eigenvalue is established. Because potential functions have only 2 dimensions, a molecular case will be studied (inversion reaction of ammonia), and discussed in a separate section. The dependency has been examined on the midpoint transition state potential energy function. The results are shown in Figure 6.8 from where we infer the safe region to be $7 < N < 10$.

Hammond and Restricted Hammond Adapted LTP Models

Quapp [13] has recently presented a new TS search algorithm which has its relevant feature an emphasis on step searching. The procedure corrects its direction, after each application, by performing a downhill step. However, it uses the Hessian and needs many steps in order to locate the TS. Using Quapp's potential function:

Table 6.5. Evolution of the uphill walk of the Approximate LTP technique applied to Cerjan-Miller's potential function to test the effect of the size of the step (Case 2, see text) on the search of the TS. Displayed are the LTP cycles (i), the step factor (N_i), the geometry, energy and gradients. At and after cycle 6 the coordinates are re-labeled. The LTP found TS (midpoint between the last projected reactants and products (i = 12)), and the expected one are displayed at the end.

		REACTANTS				PRODUCTS			
i	N_i	X	Y	E	g(x,y)	X	Y	E	g(x,y)
1	10.00	0.0000	0.0000	0.0000	0.0000, 0.0000	2.7000	0.0500	0.0062	-0.0231, 0.0495
2	8.00	0.2700	0.0063	0.0678	0.4654, 0.0054	2.4302	0.0316	0.0166	-0.0649, 0.0306
3	6.17	0.5399	0.0099	0.2178	0.5715, 0.0056	2.2049	0.0206	0.0378	-0.1317, 0.0191
4	4.17	0.8098	0.0097	0.3404	0.2893, 0.0031	1.9350	0.0137	0.0886	-0.2512, 0.0113
5	2.17	1.0798	0.0097	0.3634	-0.1117, 0.0027	1.6650	0.0098	0.1734	-0.3689, 0.0064
6	2.50	0.9178	0.0097	0.3628	0.0420, 0.0027	0.9718	0.0089	0.3673	0.0420, 0.0024
7	2.50	1.0150	0.0084	0.3677	-0.0219, 0.0022	1.0367	0.0085	0.3669	-0.0529, 0.0023
8	2.50	0.9933	0.0079	0.3679	0.0099, 0.0021	0.9977	0.0079	0.3679	0.0034, 0.0021
9	2.50	1.0046	0.0074	0.3679	-0.0068, 0.0020	1.0081	0.0075	0.3678	-0.0119, 0.0020
10	2.50	1.0005	0.0071	0.3679	-0.0007, 0.0019	1.0018	0.0071	0.3679	-0.0026, 0.0019
11	2.50	0.9990	0.0082	0.3679	0.0015, 0.0022	0.9995	0.0059	0.3679	0.0007, 0.0016
12		0.9995	0.0059	0.3679	0.0007, 0.0016	1.0005	0.0071	0.3679	-0.0007, 0.0019
TS_{ltp}		1.0000	0.0065	0.3679	0.0000, 0.0017				
TS_{ex}		1.0000	0.0000	0.3679	0.0000, 0.0000				

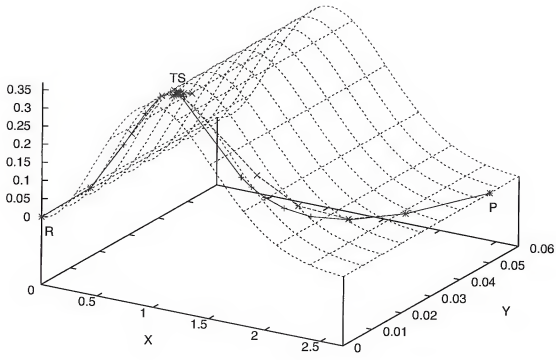


Figure 6.7. Representation of the effect of the update on the step (Case 2) in the Approximate LTP TS search in Cerjan-Miller's $E_{\text{CM}}(x,y) = (a - by^2)x^2\exp(-x^2) + \frac{c}{2}y^2$ potential energy function (we used $a = b = c = 1$). The different uphill paths followed by the classical model (continuous line), that is, N = 10 for all cycles, and the updated step factor (dashed line) techniques are plotted here, showing their convergency to the TS (bold cross) through similar, but not identical, pathways starting from the same reactants and products (R and P, respectively).

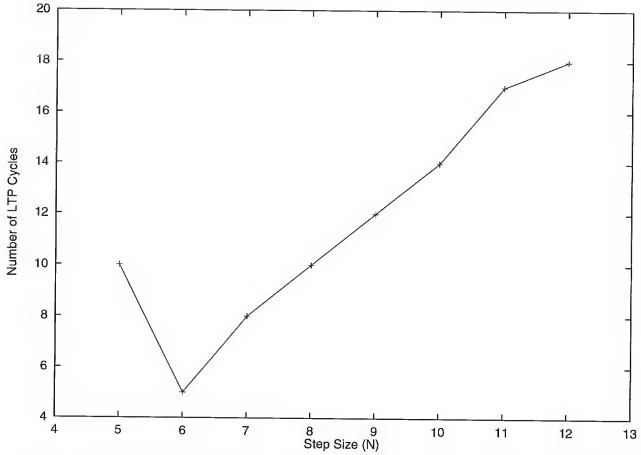


Figure 6.8. Number of Line-Then-Plane cycles needed to find the TS for the Midpoint Potential energy function: $E_{\mathrm{MPTS}}(x,y) = \left[(x-1)(x-2)\right]^2 + \left[(y-1)(y-2)\right]^2$ as a function of the Step Size.

Table 6.6. Coordinates, potential energies and number of cycles for Quapp's [13] potential energy function. A comparison with LTP, HALTP and RHALTP techniques is shown. The first and second row of all 3 LTP procedures corresponds to Exact and Approximate procedures, respectively. Calculations were performed using an step of N_o = 10 and α = 0.3. Convergency was achieved when the displacement vector norm was less than 10^{-4} at which the gradient norm was less than 10^{-4} (all units are arbitrary). The expected results for this potential energy function are those obtained by Quapp.

	X	Y	E(x,y)	Cycles
Quapp[a]	0.00	-1.00	-1.00	130
LTP[b]	0.02	-1.05	-1.00	59
	0.01	-1.06	-1.00	15
HALTP[b]	0.02	-1.05	-1.00	57
	0.02	-1.04	-1.00	13
RHALTP II[c]	0.01	-1.04	-1.00	45
	0.02	-1.04	-1.00	10

a. These results comes from reference [13]. For his method Quapp used second derivatives.

b. Deviations for both, exact and approximate LTP methods are small and very close among each other. LTP founds the TS to be located in the Reactants region.

c. Deviations for the exact and approximate of the Products Restricted Hammond Adapted LTP (RHALTP II) method are small and very close to each other.

$$E_{Quapp}(x,y) = 2y + y^2 + (y + 0.4x^2)\, x^2 \qquad (145)$$

the difference between the 3 main LTP techniques was also studied, namely LTP, HALTP and the Restricted HALTP (RHALTP) procedures. The results are shown in Table 6.6.

The function in question has a saddle point at (0.00,-1.00). For our study, we have used the points (1.77, -2.55) and (-1.00, -1.00) as reactants and products, respectively.

In Figure 6.9 we represent the behavior of both LTP techniques in this potential energy surface.

Note that all three LTP procedures converge to a very reasonable saddle point in less than half the number of the steps required by Quapp's method. The Approximate LTP methods are particularly interesting as they give the right answer with a relatively small amount of computational effort.

Summary of Results

Most methods that search for TS require an accurate evaluation of the Hessian as they proceed uphill from product to reactant, or from both points to the saddle point. These evaluations of the Hessian are costly in computer time and in storage. The LTP methods described here do not need the accurate calculation of the Hessian, with the exception of the last step, at which the Hessian should be calculated in order to check its signature. This distinction potentially allows the study of much larger chemical systems. The procedure is characterized by consecutive uphill climbs from reactant to product and vice-versa, with simultaneous minimization in all perpendicular directions at each step. Because of the nature of the steps, a TS will never be either missed nor passed provided that the steps are small enough. It can be shown that, because of the Newton-Raphson nature of the step that we consider [8], the LTP procedure will converge to the TS on any continuous surface. Connecting all points $\mathbf{P}_t$ and $\mathbf{R}_i$, will give an approximation to the reaction path.

On the other hand, the results displayed on Table 6.2 show that the Approximate LTP procedure (No Hessian) is as accurate as the Exact LTP procedure (BFGS technique used to up-date the Hessian) in order to find the TS, although connecting the points may not give an accurate representation of the RC when the approximate method is used.

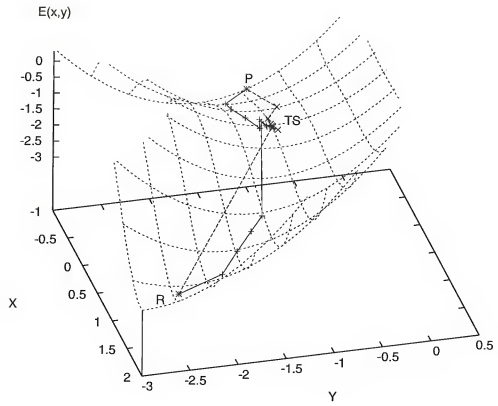


Figure 6.9. Representation of Quapp's potential energy function: $E_{Quapp}(x,y) = 2y + y^2 + (y + 0.4x^2)\, x^2$. The different uphill paths followed by the Exact LTP (continuous line) and the Exact RHALTP II (dashed line) techniques are plotted here, showing their convergency to the TS (bold cross). The plot has been deliberately tilted so the behaviour of both techniques could be better appreciated.

Finally, the testing of variations on the stepping factors has not shown to yield any improvement, or savings, in the number of cycles needed to find the TS. Consequently, it will be assumed that a step factor of 1/10 is as good as any other one, and it will be used as default for all the molecular cases studied in the next section. For each molecular system studied, the Hessian of the found TS was examined to ensure that it has one, and only one, negative eigenvalue.

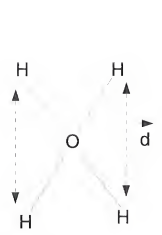
Molecular Cases for Transition State

Introduction

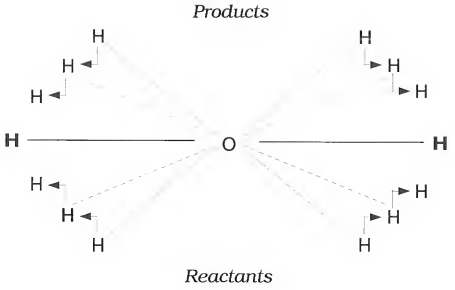
The efficiencies (accuracies) of the LTP procedures are examined by means of application to several chemical reactions in this section. We first focus on the ability of the algorithms to find the TS in some selected chemical reactions by means of the application of the LTP techniques presented in the previous chapter. All the calculations were performed at the Restricted Hartree-Fock (RHF) model, except for the thermal retro [2+2] cycloaddition reaction of Oxetane, for which we used the Unrestricted Hartree-Fock (UHF) technique so to account for breaking-formation of bonds. Calculations to test the performance of the procedure as a function of the step size then are presented.

Inversion of Water

In Figure 6.10 a scheme of the inversion of water is shown. The search direction (*d*) and the evolution of Rs and Ps towards the linear TS are shown. Whereas the hydrogen-oxygen bond length gets shortened only by 0.0207 Angstroms, the reaction is characterized by an evolution of the initial 106.81° hydrogen-oxygen-hydrogen angle to a final one of 179.67°, which is not 180° as expected (although very close) because of the LTP termination conditions, that is, the size of the displacement between the last 2 projected structures. A detailed LTP cycle-by-cycle uphill walk is shown in Table 6.7.



(a)



(b)

Figure 6.10. Symmetric inversion of water. (a) The search direction (d) is shown. b) Pictorial representation of the uphill walk of reactants and products towards the linear (bold) TS structure (not to scale).

Table 6.7. Cycle-by-cycle uphill walk in the search of the TS for the inversion of H_2O using the Exact LTP technique. Energy, geometry, the distance (square of the norm of the displacement vector), the maximum component of the gradient and number of update iterations (UI) required are displayed. The units are Hartrees for the energy, Angstroms for the oxygen-hydrogen bond length (r) and the distance, degrees for the hydrogen-oxygen-hydrogen angle (θ) and Hartrees/Angstroms for the gradient component (g_{max}). The TS is the linear structure for which $\theta = 180°$. The convergence criteria was for the maximum component of the gradient (g_{max}) of the step (in the line coordinates connecting R_5 and P_5) to be smaller than 10^{-3} (Hartrees/Angstroms). The energy and geometry of the steps in the line connecting projected coordinates are not included, but the total number of energy evaluations was 41.

	Energy	r_{OH}	θ_{HOH}	d_i	g_{max}	UI
$\mathbf{R_o = P_o}$	-18.122165	0.9952	106.74°	1.4975	0.000000	0
$\mathbf{R_1 = P_1}$	-18.115333	0.9844	122.29°	1.1980	0.000501	3
$\mathbf{R_2 = P_2}$	-18.102663	0.9798	134.35°	0.9584	0.000170	3
$\mathbf{R_3 = P_3}$	-18.090434	0.9775	143.76°	0.7667	0.000049	3
$\mathbf{R_4 = P_4}$	-18.060540	0.9745	172.85°	0.1533	0.000208	3
$\mathbf{R_5 = P_5}$[a]	-18.059076	0.9745	178.57°	0.0307	0.000010	2
$\mathbf{R_6}$ step	-18.059017	0.9744	179.71°		0.000193	0

a. At this LTP cycle, and as the distance d_5 (the square of the norm of the displacement vector) is smaller than 0.1, a threshold prestablished by the LTP algorithm, the step is amplified by a factor of 2.

The uphill evolution of the LTP search technique towards the linear TS structure of water is shown in a 3–dimensional plot in Figure 6.11

Finally, the symmetry change when going from the initial C_{2v} reactants (R) and products (P) to the $D_{\infty h}$ TS, is shown in Figure 6.12. Note that the TS has a higher symmetry than R's and P's. We shall see that this seems to be the case for molecular systems when R's and P's are mirror images of one another.

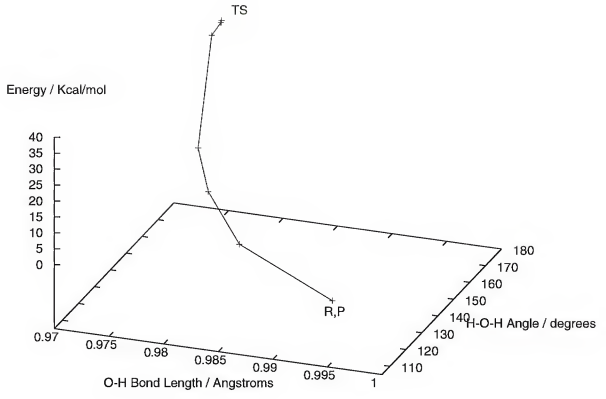


Figure 6.11. The uphill path towards the TS, according to the Exact LTP technique, starting from reactants or products (R,P), for the water inversion reaction.

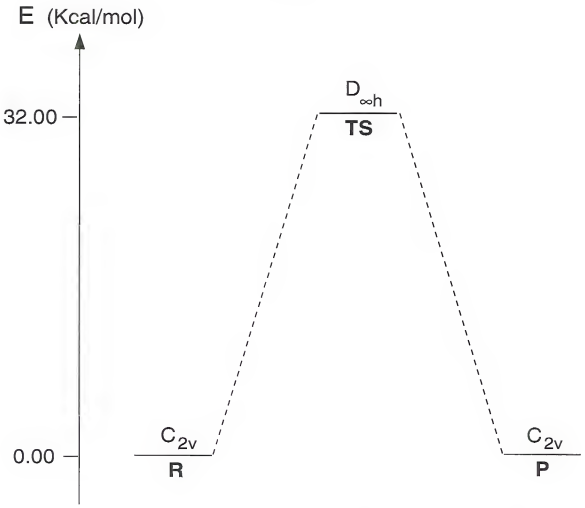


Figure 6.12. Energy and symmetry change between the initial reactants (R), products (P) and the transition state (TS) found for the inversion of water.

Symmetric Inversion of Ammonia (NH_3)

Figure 6.13 shows a 3–dimensional scheme for the inversion of ammonia, including both the initial reactants and products geometries as well as the TS found. A detailed LTP cycle-by-cycle uphill walk which includes energy, geometry and the number of iterations used to update the Hessian is summarized in Table 6.8, and graphically shown in Figure 6.14.

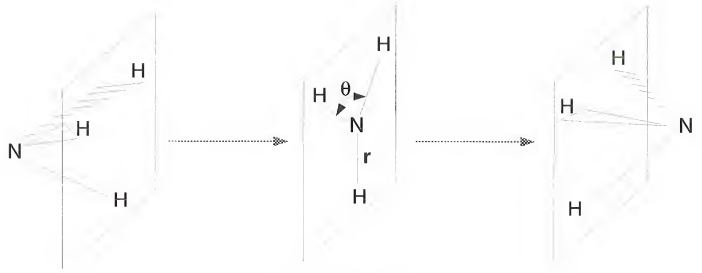


Figure 6.13. Symmetric inversion reaction of Ammonia. The initial reactants, products and the TS structure found are shown. r is the nitrogen-hydrogen bond length and θ is the hydrogen-nitrogen-hydrogen angle.

Table 6.8. LTP geometry, energy and number of iterations required to find the TS for the inversion of NH_3. The results are from the Exact Line-Then-Plane technique. The energy (Hartrees), the optimized geometry in the plane perpendicular to the walk direction, namely the nitrogen-hydrogen bond length (r, in Angstroms) and the hydrogen-nitrogen-hydrogen angle (θ, in degrees), the distance (angstroms) between R's and P's, the maximum component of the gradient and the number of update iterations (UI) of the Hessian at each LTP cycle are displayed. The convergence criteria was that the maximum component of the gradient and the distance to be smaller than 10^{-3} (Hartrees/Angstroms). The total number of energy evaluations was 31.

	Energy	r_{NH}	θ_{HNH}	d_i	g_{max}	UI
$\mathbf{R_0 = P_0}$	-12.522855	1.0503	105.63°	1.1825	0.000000	0
$\mathbf{R_1 = P_1}$	-12.520988	1.0452	110.54°	0.9460	0.000059	3
$\mathbf{R_2 = P_2}$[a]	-12.517629	1.0421	113.83°	0.7568[a]	0.000627	2
$\mathbf{R_3 = P_3}$	-12.507616	1.0383	119.74°	0.1514	0.000206	3
$\mathbf{R_4 = P_4}$	-12.507098	1.0381	119.99°	0.0303[b]	0.000480	2
$\mathbf{R_5}$ step	-12.507077	1.0381	120.00°		0.000480	0

a. One of the acceleration flags of LTP is turned on such that the step factor now is amplified by 4.

b. At this LTP cycle, and as the distance d_4 (the square of the norm of the displacement vector) is smaller than 0.1, a threshold prestablished by the LTP algorithm, the step is amplified by a factor of 2.

Our results are collected and compared with those belonging to the Augmented Hessian (AH) technique, in Table 6.9. They show that the TS found by LTP is as good as the one found using the AH method. The main difference between both procedures is that AH uses analytical second derivatives and better represents the PES using 7 SCF iterations (see Table 6.9), whereas LTP updates the Hessian using the BFGS [21] technique (other techniques as Murtagh and Sargent (MS) [20], Davidon, Fletcher and Powell (DFP) [22], and Greenstadt [23] are also available), and walks uphill towards the TS from both sides, reactants and products of the reaction.

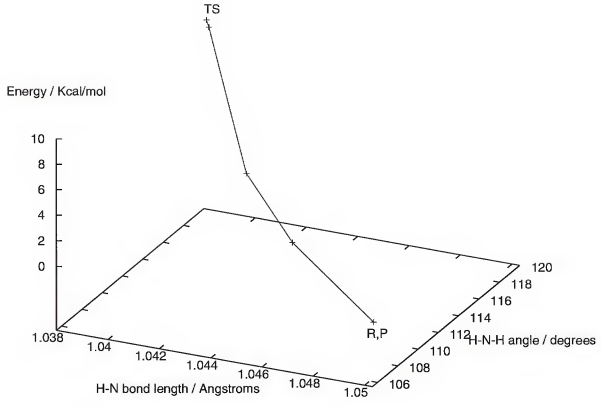


Figure 6.14. The uphill path towards the TS starting from reactants or products (R,P), for the symmetric ammonia inversion reaction.

Table 6.9. LTP geometry, energy and number of iterations required to find the TS for the Symmetric NH_3 isomerization reaction. We show results obtained using the Augmented Hessian (AH) technique, the Exact and Approximate Line-Then-Plane (LTP) techniques and a final geometry optimization (GOPT) calculation performed on the TS found by the Approximate LTP procedure. r represents the nitrogen-hydrogen bond length (in Angstroms), θ is the hydrogen-nitrogen-hydrogen angle (in degrees) and E is the energy (in Hartrees). Finally the number of SCF cycles used by each procedure are also displayed. The convergence criteria was for both the norm of the gradient and the displacement vector to be smaller than 10^{-3}.

	AH	Exact-LTP	App-LTP	App-LTP GOPT
r	1.0382	1.0382	1.0004	1.0380
θ	120°	120°	120°	120°
E	-12.507084	-12.507079	-12.498321	-12.507076
Cycles	7	22	10	3

The importance of the last condition was already described and discussed in the previous chapter. On the other hand, the Approximate LTP obtains a good answer without losing much in accuracy, using a number of SCF cycles (10) markedly lower than the one used by LTP (22). However, the App-LTP has an intrinsic problem, it does not contain enough information about the topography of the PES. Consequently it requires a final geometry optimization of the TS. When this last calculation is performed, only 3 extra SCF cycles are required. A substantial improvement is achieved in this way, as is shown in Table 6.9. Performing this final optimization is not risky at all because the TS found by App-LTP has the right signature over the Hessian as it is in the surroundings of the TS.

Finally, and for a better description of this reaction, the symmetry change when

going from the initial reactants (R) and products (P) to the TS is shown in Figure 6.15. It is interesting to note that the TS has a higher symmetry than R's and P's. This was also the case for the invertion reaction of H_2O. We shall see that this will not be the case for molecular systems where R's and P's are a mirror image of one another.

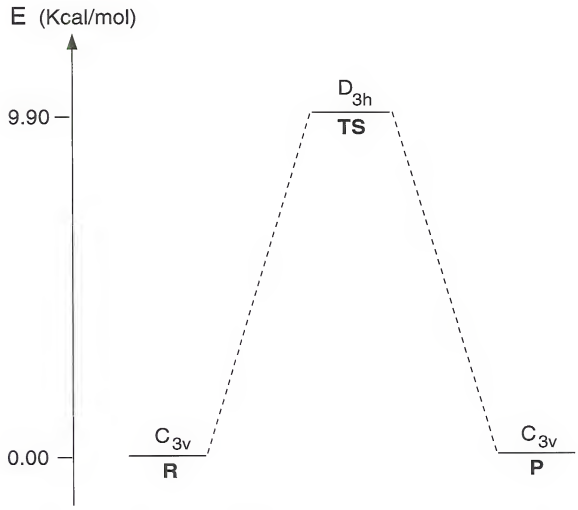


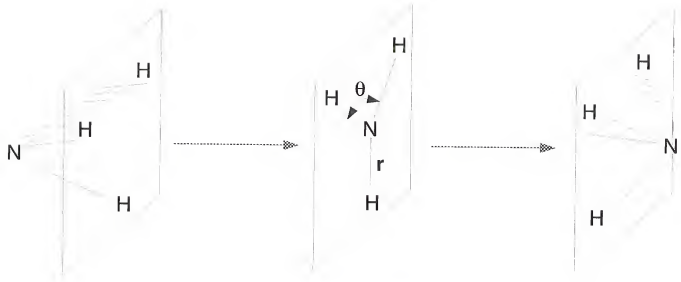
Figure 6.15. Schematic representation of the change in energy and symmetry between the initial reactants (R) and products (P), and the found transition state (TS) for the inversion reaction of ammonia.

Asymmetric Inversion of Ammonia (NH_3)

Figure 6.16 shows the geometry for the initial reactants and products as well as for the TS found. The results are displayed in Table 6.10. It is noticeable that the TS found from LTP is as good as the one resulting from the application of the AH method. The differences between both procedures are the same as described in the previous example. The same comments about the number of SCF calculations, the final geometry optimization and the symmetry changes between R's, P's and TS are valid.

Table 6.10. LTP geometry, energy and number of iterations required to find the TS for the Asymmetric NH_3 Isomerization reaction. We show results obtained using the Augmented Hessian (AH) technique, the Exact and Approximate Line-Then-Plane (LTP) techniques and a final geometry optimization (GOPT) calculation performed on the TS found by the Approximate LTP procedure. r represents the nitrogen-hydrogen bond length (in Angstroms), θ is the hydrogen-nitrogen-hydrogen angle (in degrees), E is the energy (in Hartrees). Finally the number of SCF cycles used by each procedure are also displayed. The convergence criteria was for both the norm of the gradient and the displacement vector to be smaller than 10^{-3}.

	AH	Exact-LTP	App-LTP	App-LTP-GOPT
r	1.0382	1.0380	1.0367	1.0381
θ	120°	120°	120°	120°
E	-12.507084	-12.507079	-12.507066	-12.507080
Cycles	7	22	10	2



Reactants
r = 1.0501 Å
θ = 105.70°

Transition State
r = 1.0380 Å
θ = 120°

Products
r = 0.9706 Å
θ = 119.16°

Figure 6.16. Asymmetric inversion reaction of Ammonia. The initial reactants and products, and the found transition state structures are shown. Products are energetically higher than reactants. Notice that products are a flattened mirror image of reactants.

Rotated Symmetric Inversion of Ammonia (NH_3)

This test was performed to establish whether the LTP procedure is capable of reaching the well-known planar TS conformation of NH_3 by means of a rotation through 50 degrees of the products (inverted umbrella structure), relative to the reactants (umbrella conformation) in the plane that contains the hydrogen atoms (Figure 6.17), by means of a step-by-step procedure during the up-hill walk.

These results displayed in Table 6.11 show that with only a few more SCF calculations, LTP (exact) is not only able to rotate the reactants and products towards the TS, but also to eliminate the problem associated with translations and rotations. On the other hand, the TS found by the approximate technique (Table 6.11) gives a quite

Table 6.11. LTP geometry, energy and number of iterations required to find the TS for the Rotated Symmetric NH_3 Isomerization reaction. We show results obtained using the Augmented Hessian technique, the exact and approximate Line-Then-Plane (LTP) techniques and a final geometry optimization (GOPT) calculation performed on the TS found by the Approximate LTP procedure. **r** represents the nitrogen-hydrogen bond length (in Angstroms), θ is the hydrogen-nitrogen-hydrogen angle (in degrees), **E** is the energy (in Hartrees). Finally the number of SCF Iterations (energy evaluations) used by each procedure are also displayed. The convergence criteria was for both the norm of the gradient and the displacement vector to be smaller than 10^{-3}.

	AH	Exact-LTP	App-LTP	App-LTP-GOPT
r	1.0382	1.0381	1.0370	1.0381
θ	120°	120°	120°	120°
E	-12.507084	-12.507079	-12.507070	-12.507080
SCF-Its	7	26	12	2

good answer, and its optimization (App-LTP-GOPT), which requires only two more energy evaluations gives a final answer as good as does the AH technique. This outcome is in agreement with the cases studied before.

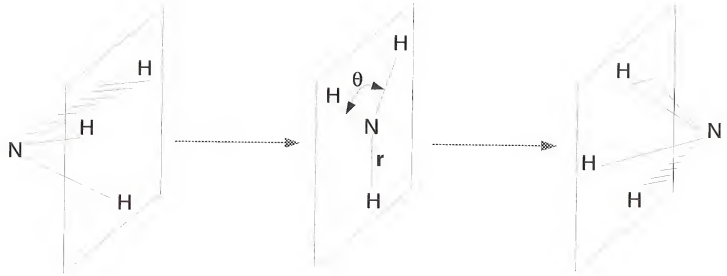
Hydrogen Cyanide: HCN → CNH

Figure 6.18 shows a scheme of the HCN proton transfer reaction. The initial reactant and product and the TS geometries as well as a detailed LTP cycle-by-cycle uphill walk which includes energy, geometry, and the number of update iterations to update the Hessian is summarized in Table 6.12.

The LTP uphill walk towards the transition state (TS) starting from reactants (R) is represented in Figure 6.19. In spite of the tortuous shape of the hypersurface the algorithm is capable of overcoming it.

A comparison of our results on the search for the TS of the proton transfer type reaction with those of other research group's collected in Table 6.13. Zerner et. al. have used the augmented Hessian (AH) [27] technique, but no bond lengths were reported. Bell and Crighton [87] on the other hand, have used an INDO Hamiltonian for their study. However, their results cannot be compared with ours derived from the application of ZINDO, because both methodologies do not have the same parametrization. Their results are included here for geometry comparison purposes only.

In relation to the geometry the angular coordinate obtained by Zerner et.al. [18] is reproduced, the difference being only 0.09°. There is, however, a difference of 0.038 A for the H-C bond and of 3.86° for the HCN angle between our data and those of Bell and Crighton, which we associate with the different parametrization of the Hamiltonians. The energy the difference between AH and this work is 0.3 Kcal/mol.



Reactants
r = 1.0508 Å
θ = 105.59°

Transition State
r = 1.0381Å
θ = 120.00°

Products
r = 1.0508 Å
θ = 105.59°

Figure 6.17. Rotated symmetric inversion reaction of Ammonia. The initial reactants and products, and the found transition state geometries are displayed. Products are a mirror image of the reactants but rotated by 50 degrees, but are geometrically and energetically identical.

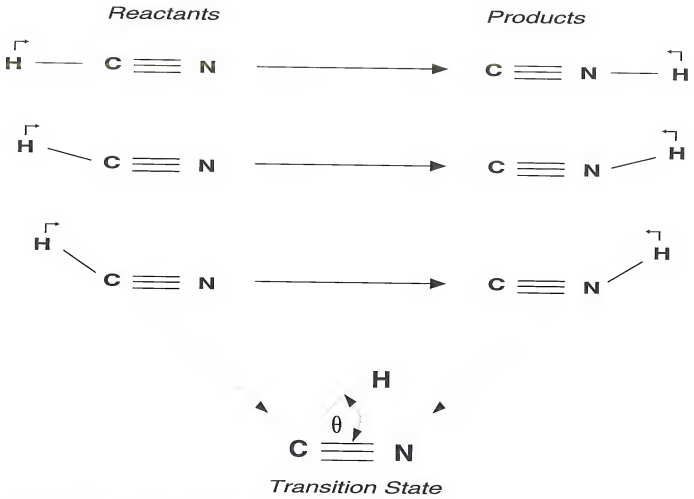


Figure 6.18. Proton transfer reaction of Hydrogen Cyanide. The behavior of the LTP technique is shown as reactants and products suffers structural modifications (arrows), to coalesce finally in the TS.

Table 6.12. LTP geometry, energy and number of update iterations (UI) required to find the TS for the hydrogen cyanide isomerization reaction (HCN —> CNH). The results are those belonging to the Exact Line-Then-Plane technique. The optimized geometry in a plane perpendicular to the walk direction, namely the carbon-nitrogen and hydrogen-carbon bond lengths (in Angstroms) and the hydrogen-carbon-nitrogen angle (θ, in degrees), the energy (Hartrees) at each iteration and the number of update iterations of the Hessian for each cycle are displayed. The convergence criterion requires that the maximum component of the gradient should be smaller than 10^{-3} (Hartrees/Angstroms), whereas the search was started at an angle of θ = 130°. The energy and geometry of the steps in the line are not shown, but the total number of energy evaluations was 32.

	Energy	r_{HC}	r_{CN}	r_{HN}	θ_{HCN}	θ_{HNC}	UI
R_o	-17.354816	1.0752	1.1840		130.00°		0
P_o	-17.313130		1.1991	1.0376		180.00°	0
R_1	-17.304199	1.0766	1.2161		104.65°		4
P_1	-17.312384		1.2024	1.0521		175.57°	4
R_2	-17.283882	1.1226	1.2285		77.78°		4
P_2	-17.291977		1.2041	1.0074		144.09°	3
R_3	-17.283736	1.0953	1.2195		88.72°		3
R_4	-17.282221	1.1093	1.2244		82.10°		2
P_4	-17.282171	1.1039	1.2232		84.31°		2
R_5	-17.282200	1.1092	1.2247		82.23°		2
P_5	-17.282156	1.1043	1.2233		84.16°		2

In order to obtain further insight on the origin of these differences (that are, however, very small), we have performed a geometry optimization of the TS found by LTP, which required only 2 more SCF iterations. As a result, the geometry has improved, but not significantly. Whereas the difference in the HC bond length decreased only 0.005 Angstroms, the CN bond length increased in 0.002 Angstroms and the HCN angle by

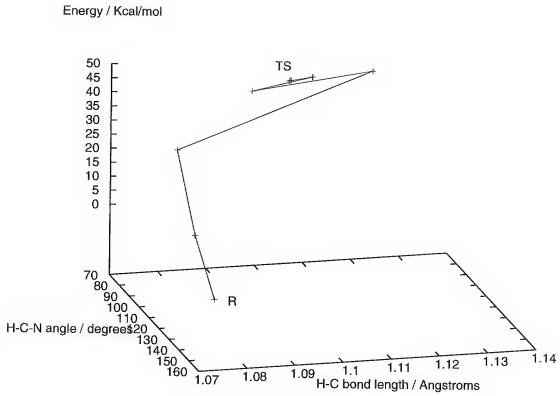


Figure 6.19. The uphill path towards the TS, starting from reactants (R), for the isomerization reaction of hydrogen cyanide.

0.22°. On the other hand, the decrease in energy accounts to only 0.3 Kcal/mol. The number of iterations was not comparable because all the calculations have different starting geometries for the HCN angle α_0 as shown in Table 6.12. However, it is interesting to notice that the number of SCF calculations was still in the range of procedures associated with the use of second derivatives.

Table 6.13. Geometry, energy and number of SCF iterations required for the convergence to transition state for the HCN isomerization reaction. Displayed are the results obtained through the Augmented Hessian technique, Bell-Crighton and LTP techniques. Also displayed in the last column is the TS found when an extra geometry optimization calculation was performed on the LTP transition state already found. r_{CN} and r_{HC} are the corresponding carbon-nitrogen and hydrogen-carbon bond lengths (Angstroms), θ is the hydrogen-carbon-nitrogen angle (in degrees) and E is the energy (in Hartrees).

	AH[a]	Bell-Crighton[b]	LTP[c]	GOPT on LTP
r_{CN}		1.2254	1.2255	1.2239
r_{HC}		1.1397	1.1019	1.1073
θ	83°	79.23°	83.09°	83.22°
E	-17.282563		-17.282074	-17.282125
Its	15	11	13	2

a. For this technique Zerner *et.al.* [18] did not report the geometry of the TS found. Their starting hydrogen-carbon-nitrogen angle was of 180°.

b. Bell and Crighton used an INDO Hamiltonian that is different from the one used by Zerner and his coworkers, as well as by LTP. However we included it here in order to compare the geometries. Their starting hydrogen-carbon-nitrogen angle was 90°. They did not report the energetics of their calculations.

c. The starting hydrogen-carbon-nitrogen angle was of θ = 160°.

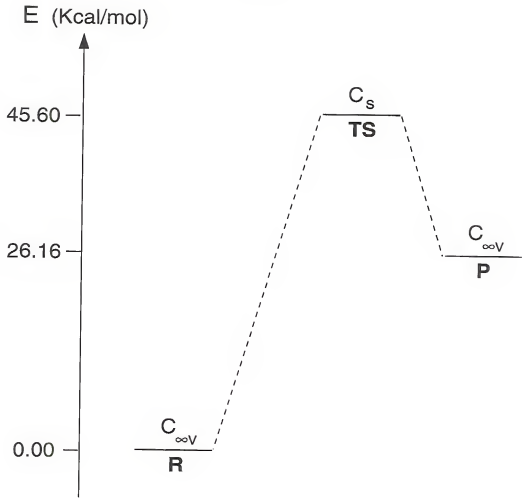


Figure 6.20. Schematic representation of the change in energy and symmetry between the initial reactants (R) and products (P), and the found TS for the isomerization of hydrogen cyanide.

From Table 6.12 we notice that the TS is located by the second iteration. It is in the reactants zone, with an HCN angle of 77.78°. From this point, the LTP algorithm resets the coordinates and searches in a smaller part of the hypersurface, accelerating convergence. At the next iteration (3) a test is performed at the reactants zone. The TS is again located in a narrower region of the space, and once more the coordinates are reset accordingly. A few final iterations are performed and the TS is found.

The examination of the change of symmetry in this reaction shows that, in contrast to the previous cases, the TS has lower symmetry (C_s) than either the reactants or the products ($C_{\infty v}$) as shown in Figure 6.19.

Formic Acid

In Figure 6.21 the [1,3] Sigmatropic reaction of Formic Acid under study is shown. The reaction is characterized by the migration of the hydrogen with its sigma bond in a π carboxilic, *i.e.* the migration occurs by a shift in the π bonds of this metanoic acid environment.

The energetical and geometrical evolution of the LTP TS search technique for this reaction is shown in Table 6.14. The TS is characterized by the migrating proton bonded to both, the source and migration, oxygens, by a oxygen-hydrogen bond length of r_{24} = 1.2045 Angstroms. Both oxygens form angles with the carbon and the hydrogen atached to the carbon of $\theta_{215} = \theta_{315} = 129.82°$. As expected, both oxygen carbon bonds have now the same length, $r_{12} = r_{13}$ = 1.2960 Angstroms, and the oxygen-carbon-oxygen angle has narrowed by around 23°.

It is assumed that the mechanism of the reaction is a Symmetry-allowed Suprafacial sigmatropic shift reaction as reactants, products and TS are all in the same molecular plane.

The LTP uphill search towards the TS starting from reactants or products (R,P) is represented in Figure 6.22, where the energy has been plotted against the most relevant geometrical changes during the simulated reaction.

These geometrical rearrengements convey some changes in the symmetry of the main structures of this reaction. The TS goes to a lower C_{2v} symmetry in relation to the $C_{\infty v}$ of both reactants and products, as shown in Figure 6.23.

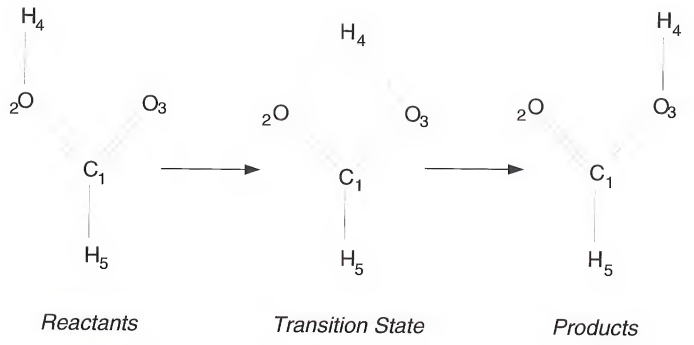


Figure 6.21. The [1,3] Sigmatropic reaction of formic acid. The initial reactants, the transition state found and initial products structures are shown. Products are a mirror image of reactants. The geometrical details are given in Table 6.13.

Table 6.14. LTP geometry, energy and number of update iterations (UI) required to find the TS for the formic acid [1,3] Sigmatropic reaction. The results are those belonging to the exact Line-Then-Plane technique. As reactants (R) and products (P) had the same geometry and energy we group them as: $\mathbf{Q_i = R_i = P_i}$, for i = 0, 1, 2, 3, 4 . The optimized geometry (bond lengths **r** are in Angstroms and angles θ are in degrees) in a plane perpendicular to the walk direction at each step, the energy (Kcal/mol) and the number of update iterations of the Hessian for each cycle are displayed. The convergence criterion requires that the maximum component of the gradient should be smaller than 10^{-3}. The energy and geometry of the steps in the line are not shown, but the total number of energy evaluations was 37. The labeling of the atoms is as shown in Figure 6.20.

	Energy[a]	$\mathbf{r}_{12}$	$\mathbf{r}_{13}$	$\mathbf{r}_{24}$	θ_{213}	θ_{215}	θ_{315}	θ_{124}	UI
$\mathbf{Q_0}$	0.0000	1.3298	1.2451	0.9976	123.15°	112.75°	124.10°	113.17°	0
$\mathbf{Q_1}$	15.1640	1.3304	1.2527	1.0238	110.01	123.37°	126.63°	89.93°	4
$\mathbf{Q_2}$	32.7422	1.3157	1.2883	1.0764	104.82°	128.68°	126.50°	82.72°	3
$\mathbf{Q_3}$	38.6872	1.3414	1.2421	1.1181	105.33°	126.19°	128.48°	79.18°	3
$\mathbf{Q_4}$	47.7016	1.3083	1.2917	1.1656	100.55°	130.20°	129.26°	76.60°	2
$\mathbf{R_5}$	49.2611	1.2960	1.2960	1.2045	100.37°	129.82°	129.82°	74.08°	2

a. The origin of the energy (at Q_0) is: E = –41.394497 a.u.

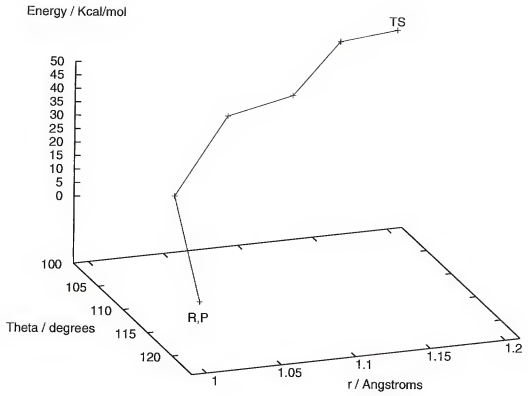


Figure 6.22. The uphill path towards the transition state (TS), starting from reactants (R), for the [1,3] Sigmatropic reaction of formic acid.

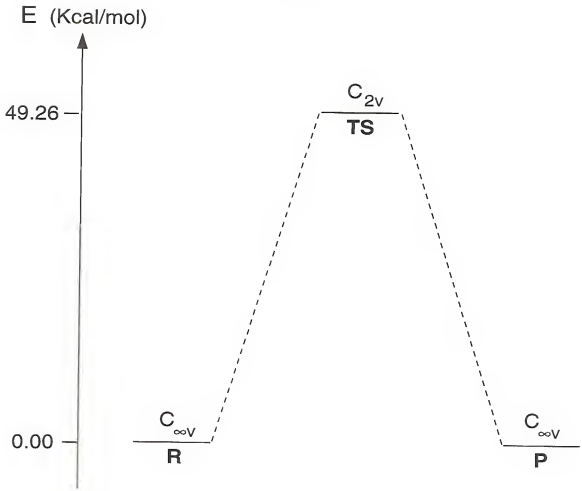


Figure 6.23. Schematic representation of the change in energy and symmetry between the initial reactants (R) and products (P), and the transition state (TS) found for the Sigmatropic shift reaction of formic acid.

Methyl Imine

This is an interesting isomerization reaction that occurs through 2 main mechanisms as shown in Figure 6.24.

The first mechanism (a) involves the move, in the molecular plane, of the hydrogen (H_5) attached to the nitrogen passing through a planar C_{2v}. The results for the LTP search for this mechanism are shown in Table 6.15, indicating that the TS is found at

the 7th cycle, in the reactants step with a hydrogen-nitrogen-carbon angle θ_{521} closer to 180° as expected. The other significative change in the geometry is that the hydrogen-nitrogen bond length is shortened during the reaction as disclosed from Table 6.15.

The second mechanism (b) involves the internal rotation of the imine double bond. This mechanism is performed through a reaction coordinate chosen to be the dihedral angle α, as shown in Figure 6.24. This mechanisms occurs when the hydrogen connected to the nitrogen (H_5) comes out of the molecular plane passing through a maxima at α = 90°. The results for this mechanism have been fitted through an interpolation procedure [5] and are represented here in terms of the dihedral angle α, which is used as the reaction coordinate. The results for this mechanism are shown in Table 6.16, in which only the most relevant geometrical changes have been displayed.

Table 6.16. Internal rotation of methyl imine according to mechanism (b) of Figure 6.23. The TS is founded at a dihedral angle of α = 90°. The labeling of the geometrical parameters is as shown in Figure 6.23. Notice that this internal reaction is symmetric.

α	Energy[a]	r_{25}	θ_{125}
0° = 180°	0.0000	1.0546	112.72°
20° = 160°	8.8409	1.1026	111.74°
40° = 140°	31.2268	1.2242	109.27°
60° = 120°	56.6831	1.3624	106.46°
80° = 100°	73.2985	1.4526	104.62°
90°	75.5774	1.4650	104.37°

a. The origin of the energy (at P_0) is: E = –40.490650 a.u.

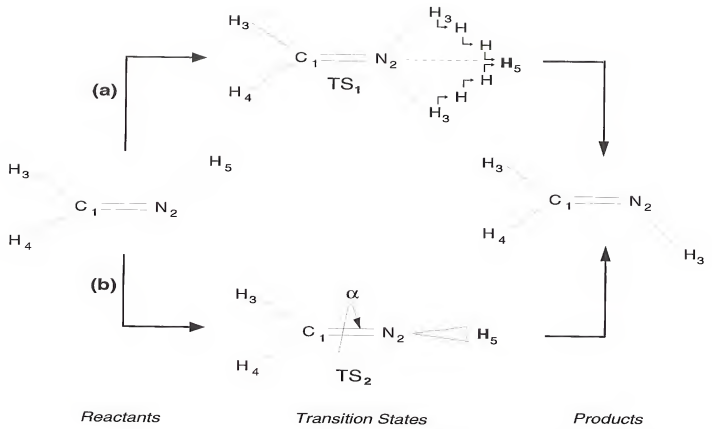


Figure 6.24. The methyl imine isomerization reaction. The initial reactants, the 2 possible transition states and the initial products structures are shown for the 2 pathways for this reaction. a) Transition state (TS_1) found by the Exact LTP technique, showing its uphill behaviour. b) A second transition state (TS_1) due to the internal rotation of the double bond. Products are a mirror image of reactants. The geometrical details are given in Table 6.15.

Table 6.15. LTP geometry, energy, and number of update iterations (UI) required to find the TS for the methyl imine isomerization reaction, according to mechanism (a) of Figure 6.23. The results are those belonging to the exact Line-Then-Plane technique. As reactants (R) and products (P) had the same energy and almost the same geometry, we group them as: $\mathbf{Q_i} = \mathbf{R}_i = \mathbf{P_i}$, for i = 0, 1, 2, 3, 4. The optimized geometry (bond lengths r are in Angstroms and angles θ are in degrees) in a plane perpendicular to the walk direction at each step, the energy (Kcal/mol) and the number of update iterations of the Hessian for each cycle are displayed. The convergence criterion requires that the maximum component of the gradient should be smaller than 10^{-3}. The energy and geometry of the steps in the line are not shown, but the total number of LTP cycles was 73. The labeling of the atoms is as shown in Figure 6.23.

	Energy[a]	r_{12}	r_{13}	r_{14}	r_{25}	θ_{213}	θ_{214}	θ_{314}	θ_{125}	UI
$\mathbf{Q_0}$	0.0000	1.2884	1.0941	1.0941	1.0546	122.38°	122.38°	115.24°	112.72°	0
$\mathbf{Q_1}$	5.5596	1.2812	1.0870	1.1019	1.0323	121.79°	123.02°	115.20°	131.08°	6
$\mathbf{Q_2}$	12.6725	1.2766	1.1025	1.0883	1.0291	123.07°	121.90°	115.04°	142.77°	4
$\mathbf{Q_3}$	18.1732	1.2737	1.1029	1.0895	1.0289	123.09°	122.00°	114.91°	151.05°	5
$\mathbf{Q_4}$	22.0346	1.2714	1.1027	1.0910	1.0291	123.08°	122.13°	114.80°	157.22°	5
$\mathbf{Q_5}$	29.1952	1.2679	1.0996	1.0969	1.0299	122.83°	122.61°	114.56°	175.56°	6
$\mathbf{Q^b_6}$	29.4964	1.2679	1.0983	1.0978	1.0298	122.73°	122.68°	114.59°	179.11°	3
$\mathbf{R^c_7}$	29.5083	1.2679	1.0981	1.0980	1.0297	122.71°	122.70°	114.59°	179.82°	0

a. The origin of the energy (at Q_o) is: E = −18.888477 a.u.

b. At this cycle the step factor is reseted to 2.5, according to the LTP algorithm.

c. At this cycle the step factor is amplified by a factor of 2, according to the LTP algorithm.

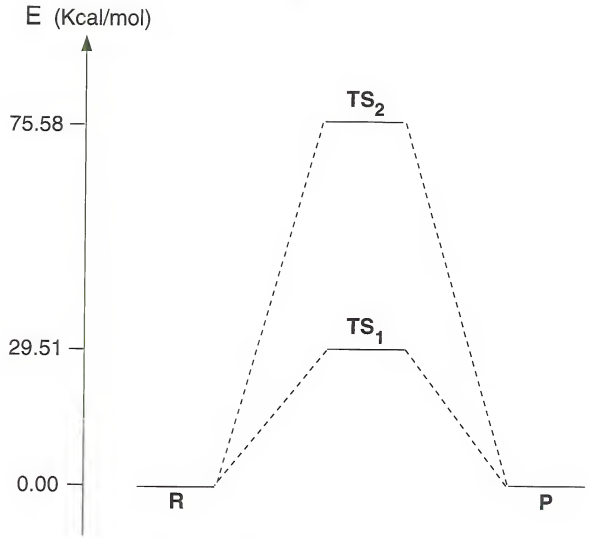


Figure 6.25. Relative energy differences between the two studied mechanisms for the methyl-imine reactions (not to scale). TS_1 corresponds to the in the molecular plane saddle point (mechanism (a)), whereas TS_2 corresponds to the out of the molecular plane saddle point (mechanism (b)), according to the scheme shown in Figure 6.24.

The energetical difference between this two possible mechanisms is shown in Figure 6.25, which indicates that the in the molecular plane TS (TS_1 through mechanism (a)) is lower in energy than the out of the molecular plane TS (TS_2 through mechanism (b)) by more than 46 Kcal/mol.

Thermal Retro [2+2] Cycloaddition Reaction of Oxetane

This mechanism considers two molecules reacting to give products. Particular care was taken on the preparation of the input file such that the molecular planes of reactants were not the same but parallel among them, and at a distance of 3.0 Angstroms (r_{17} = r_{28} = 3.0).

In Figure 6.26 the mechanism of the reaction is shown as well as the labeling of the atoms employed. The results of the LTP uphill search are shown in Table 6.17. The TS found shows an elongation of the ethane C_1-C_2 bond length, from 1.3238 Angstroms at the initial reactants structure to 1.4082 Angstroms at the TS, which is midway between the double and single bond of ethene and the oxo-cycle, respectively. The same tendency is observed for the C_7-O_8. This changes are expected as the new single bonds C_1-C_7 and C_2-O_8 are coming to a distance closer to a single bond, while the original double bonds migrate to intermediate single-double bonds.

Our results, obtained using the Restricted-Hartree-Fock method, allow us to conclude that the reaction occurs in a concerted fashion, that is, the formation and breaking of bonds occurs simultaneously, as no intermediate of reaction was found.

Step Size Dependency

The symmetric inversion reaction of ammonia has been used to study the variation on the number of SCF calculations that are necessary to find the TS. We have also focused our attention on the dependence of the convergence to the TS geometry with the step size as the Line-Then-Plane algorithm was stepping up-hill. Regardless the step sizes (N) taken, the accuracy of the TS found was kept as the energy, the nitrogen-hydrogen bond length and the hydrogen-nitrogen-hydrogen angle were kept constant, as reported in Table 6.18.

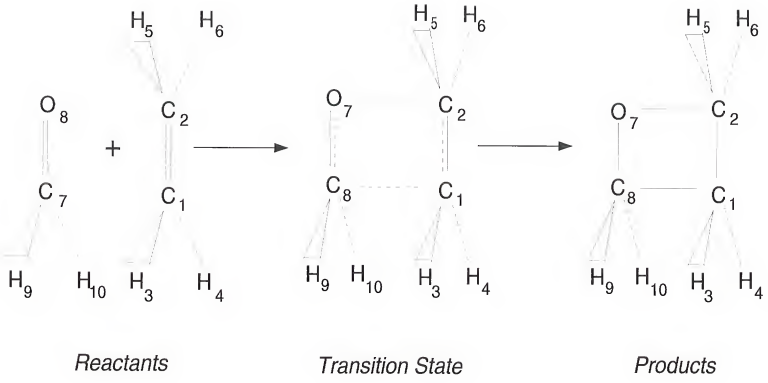


Figure 6.26. The thermal retro [2+2] cycloaddition reaction of Oxetane mechanism. The initial reactants, the transition state found and initial products structures are shown. The geometrical details are given in Table 6.17.

Table 6.17. LTP energy (Kcal/mol), geometry (bond lengths r are in Angstroms and angles θ are in degrees) and the number of update iterations (UI) required to find the TS for the thermal retro [2+2] cycloaddition reaction of Oxetane. The results belong to the Exact LTP technique. The convergence criterion was that the maximum component of the gradient and the displacement vector should be smaller than 10^{-3}. The energy and geometry of the steps in the line are not shown. The total number of LTP cycles was 47. The labeling of the atoms is as shown in Figure 6.26.

	Energy[a]	ldl	r_{12}	r_{78}	r_{17}	r_{28}	θ_{287}	UI
R_o	336.6424	3.1202	1.3238	1.2366	3.0000	3.0000		0
P_o	0.0000		1.5007	1.3849	1.5000	1.3842	88.65°	0
R_1	337.6691	2.4962	1.3403	1.25.12				2
P_1	49.9408		1.4819	1.3699	1.6498	1.5459	88.87°	2
R_2	340.6801	1.9969	1.3537	1.2629				2
P_2	135.2269		1.4671	1.3579	1.7697	1.6753	89.07°	2
R_3	344.3734	1.5976	1.3646	1.2723				2
P_3	215.0582		1.4553	1.3483		1.7788	89.23°	2
R_4	3480924	1.2781	1.3733	1.2799				2
P_4	280.1598		1.4459	1.3345				2
R_5	351.5258	1.0224	1.3804	1.2859				3
P_5	329.8571		1.4385	1.3345				3
R_6	354.5469	0.8178	1.3861	1.2907				2
P_6	364.1506		1.4325	1.3296				2
$R^b{}_7$	365.5244	0.1636	1.4045	1.3063				2
P_7	371.1769		1.4138	1.3140				2
$R^c{}_8$	367.8363		1.4082	1.3094				0

a. The origin of the energy (at P_o) is: E = –40.490650 a.u.

b. At this cycle the step factor is reseted to 2.5, according to the LTP algorithm.

c. At this cycle the TS is found in the Reactants step. LTP stops.

Table 6.18. Number of LTP iterations, SCF calculations, and variation on the energetics and structure as function of the step size (N) for the symmetric inversion reaction of ammonia. For all cases the convergence was the same as detailed above. Energy is in Hartrees, the nitrogen-hydrogen bond length is in Angstroms and the hydrogen-nitrogen-hydrogen angle (α) is in degrees.

N	LTP-Its	Iterations	Energy	r_{NH}	α
5	5	35	-12.507077	1.0380	120°
6	5	27	-12.507075	1.0376	120°
7	5	29	-12.507077	1.0381	120°
8	6	31	-12.507077	1.0380	120°
9	5	31	-12.507077	1.0380	120°
10	6	29	-12.507077	1.0381	120°
11	7	39	-12.507076	1.0378	120°
12	6	35	-12.507077	1.0381	120°

Figures 6.27 and 6.28 show the variation on the number of SCF calculations and the change of the nitrogen-hydrogen bond length as a function of the step size, respectively. The first one shows a wide safe range: $7 \leq N \leq 10$ maintain the accuracy in the TS found. On the other hand, for the nitrogen-hydrogen bond length change, it becomes clear that there is not a big error for values of N smaller than 7 and bigger than 10. However, it seems that this range is a safe one for a good answer. The same can be inferred for the number of SCF calculations required to locate TS.

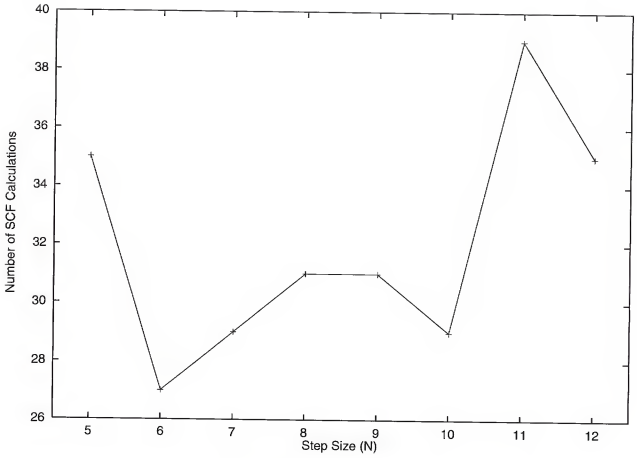


Figure 6.27. Number of SCF calculations as a function of the step size (N) for the Symmetric Ammonia Isomerization reaction.

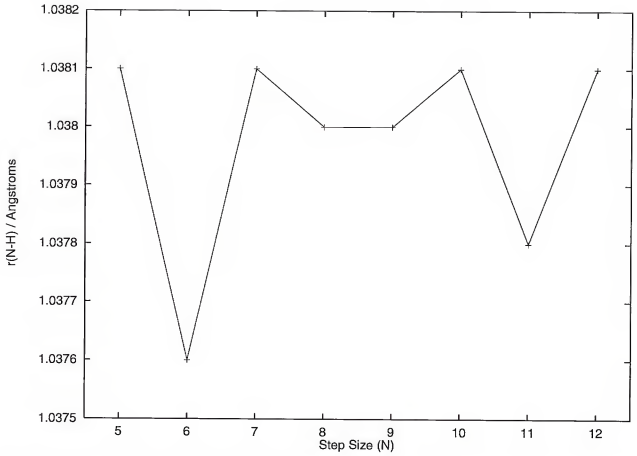


Figure 6.28. Nitrogen-hydrogen bond length (r(N-H)) change as a function of the step size (N) for the Symmetric Ammonia Inversion reaction.

Hammond Adapted LTP Results

When the Hammond-adapted techniques were applied to Quapp's potential function (Table 6.6), it was clear that a considerable reduction in the effort (*i.e.* number of function evaluations), to reach the TS was achieved. In particular, the approximate-restricted-Hammond-adapted LTP technique had a considerable reduction (2/3) in the number of energy evaluations, while maintaining the same accuracy as the exact (LTP) procedures.

In this section, the LTP technique is compared against the Hamond-adapted (HALTP), and the two restricted Hammond-adapted LTP (RHALTP I (clamped initial reactants) and RHALTP II (clamped initial products)) procedures through the study of the asymmetric NH_3 isomerization reaction. These algorithms are also compared with the augmented Hessian model. The results displayed in Table 6.19 indicate that the geometry and the energy of the TS found is kept almost constant by all methods. The HALTP procedure reduces the number of SCF calculations aproaching the number required by AH. On the other hand, RHALTP I requires the same amount of SCF calculations (7) as AH does. However, RHALTP II requires less effort than any of the other techniques studied, as it requires only 6 SCF iterations.

Summary of Results

The results of the symmetric inversion of ammonia derived from the application of the Approximate-LTP method raise the question of whether it would not be cheaper to use the Approximate method and optimize the TS found instead of using the exact LTP. In Table 6.11 we see that for the TS search on the symmetric inversion of ammonia reaction, LTP needs 22 SCF iterations, while the App-LTP + GOPT requires only 13 SCF iterations. This outcome implies that around half of the effort is required by the

Approximate-LTP technique to give roughly the same answer. The same behavior is observed for the Asymmetric inversion of ammonia reaction, according to the results displayed on Table 6.11.

It seems then that a full line search optimization in the perpendicular direction is not worthwhile and that a search using an Approximate method (like App-LTP) will be close enough to the right TS, with a simultaneous substantial reduction on the computational effort. This reduction becomes even more significant for bigger systems.

This tendency has been already observed and examined by Zerner [17, 18]. The present work confirms, hence, their observations. The TS found by LTP (when no approximations are involved) generally does not require any optimization.

As shown in Tables 6.9 to 6.11, the change in the geometry and the energy of the TS located by LTP for the different types of NH_3 reactions studied, do not improve significantly after optimization (the difference in energy goes down only by approximately 10% of an already small deviation). The improvement that results from optimization requires only a few (2 or 3) extra SCF caculations.

It is interesting to note that, in the approximate methods, the main improvement is related to the substantial reduction of the number of SCF iterations necessary to find the TS. This is the purpose of designing these methods. It is also clear that, regardless of the technique that is used, there are no significant changes in any of the relevant parameters measured and displayed in Table 6.16, that is, the accuracy is kept constant. The difference between energies and bond lengths obtained with the Augmented Hessian and the LTP family of procedures are associated whith the inherent use of analytical second derivatives by the former.

From the Hammond-Adapted and the Restricted-Hammond-Adapted LTP algo-

Table 6.19. Geometry, energy, convergence criterias and number of iterations required to find the TS for the asymmetric NH_3 isomerization reaction. We show results obtained using LTP, Approximate LTP, Hammond Adapted LTP (HALTP), the two cases of the Restricted Hammond Adapted LTP (RHALTP I and RHALTP II, exact ones) and the Augmented Hessian technique. **r** represents the nitrogen-hydrogen bond length (in Angstroms), θ is the hydrogen-nitrogen-hydrogen angle (in degrees), E is the energy (in Hartrees) and $|g|$ is the norm of the gradient achived after convergence to the TS. The convergency criteria require the norm of the gradient and the displacement vector, to be smaller than 10^{-3}. Finally the number of SCF Iterations (energy evaluations) required by each procedure is also displayed.

	LTP	App-LTP	HALTP	RHALTP I [a]	RHALTP II[b]	AH
r	1.0364	1.0367	1.0364	1.0362	1.0362	1.0382
θ	120°	120°	120°	120°	120°	120°
E	-12.507062	-12.507079	-12.507062	-12.507070	-12.507075	-12.507084
|g|	$3.71*10^{-5}$	$4.51*10^{-5}$	$3.10*10^{-5}$	$5.13*10^{-5}$	$5.13*10^{-5}$	$1.00*10^{-5}$
Its	10	10	8	7	6	7

a. Reactants coordinates are kept constant (initial ones) along all the search.

b. Products coordinates are kept constant (initial ones) along all the search.

rithms whose results are shown in Tables 6.6 and 6.16, it becomes clear that these algorithms represent a substantial reduction in the number of energy evaluations necessary to find the TS involved. Moreover, in the molecular case of the asymmetric inversion reaction of ammonia, the RHALTP I is seen to be as competitive as the AH model, whereas RHALTP II is less expensive. Note that for all LTP techniques, the structure and energy of the TS found showed no appreciable deviation among each other. For all the molecular systems studied here, our calculations shows LTP to be a reliable TS search algorithm.

The same INDO Hamiltonian, basis set and optimization techniques have been used for all the procedures tested here. Furthermore, they are comparable to those procedures that employ it, as the Augmented Hessian method, in terms of speed (number of SCF Iterations), accuracy of the calculated energy and geometry of the found TS.

Model Potential Function for ARROBA

Introduction

With the purpose of using the LTP technique features for search in perpendicular directions, a procedure to optimize geometries was developed (ARROBA) that was described in the preceding chapter. The algorithm has been tested on a potential energy function and on the H_2O molecule. The dependence of its accuracy and convergence on the line search technique parameter (α). The results presented here are, although preliminary, are promising and encouraging.

Model Potential Function

The behavior of our procedure was tested by means of its application to the same model potential function that contains a minimum, proposed in the preceding section and shown in Figure 6.6. We have choosen the initial geometry to

be at the point (0,0), from which a new point was generated and the search for a minimum started. Figure 6.29 shows this potential energy surface and the behavior of the procedure. The minima are located at (+/- 0.7071,-/+0.7071) with an energy of –0.5000 . We have used, for the perpendicular search, α = 0.2 for the first cycle and 1.00 for the rest of the cycles. Both the Exact and Approximate ARROBA techniques were checked. It is interesting to notice that both procedures, represented by a cross in the contour plot, walk downhill together using the same coordinates (positions) and direction, finding the minimum in the 10^{th} iteration. The exact method uses a total of 12 energy evaluations whereas the approximate one uses only 10. The difference is due to the BFGS update technique used by the Exact method.

The comparison of Exact and Approximate methodologies seems to indicate, as was previously observed by Zerner [17, 18], that the effort of updating the Hessian (or the use of more sophisticated techniques) is not always imperative. For several cases, approximate algorithms are good enough, providing accuracy is used in the final stages of the search.

Step Size Dependence

This geometry optimization procedure has among its outstanding features the automatic generation of steps and direction. There is no constraint for the step as each new projected point is generated using the line search technique (LST). For this generation, a Newton-Raphson-like step is taken and the second derivative matrix is replaced by the projector itself.

We have studied the dependence, and accuracy, of the number of energy evaluations as a function of the LST parameter (α), which is not related to the LTP step parameter

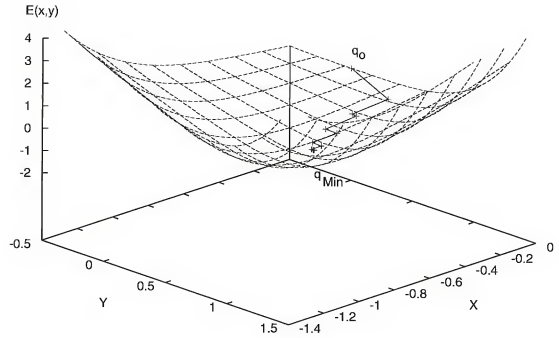


Figure 6.29. Potential energy surface with a minima $E_{\mathrm{Min}}(x, y) = x^4 + y^4 + 2xy$. The starting point is $q_o(0, 0)$ and the minimum founded at q_{Min}(–0.7071, 0.7071), with an energy of $E(q_o)$ = –0.5000. All units are arbitrary. Notice that both procedures (exact and approximate ARROBA) represented by a cross in the contour plot walk downhill together using the same coordinates and direction, finding the minimum at the 10^{th} iteration.

(N), using only the Approximate ARROBA method. Our results are displayed in Table 6.20 and graphically shown in Figure 6.29. The accepted values for the minima of this function are E = –0.5000 and (x,y) = (+/-0.7071, -/+0.7071) arbitrary units.

Summary of Results

The ARROBA procedure for geometry optimization is accurate in finding the minima for a wide range of values of alpha. The exception is defined by values of α in the vicinity of α = 0.3, as can be inferred from Table 6.16. In this range of α values, the deviation for the energy and coordinates are ΔE = 0.0036, ΔX = 0.0344 and ΔY = 0.0015. The study of the dependence of the step size on α was aimed at establishing a range of values of alpha for which not only a good response from the algorithm, in terms of both the energy and the coordinates, but also a reduced number of energy evaluations that speeds up the calculations was achieved. According to our research, a convenient, and safe, range of alpha's to consider will be: $0.18 < \alpha < 0.25$. Values of α = 0.20 and 0.21 are recomended, as they not only give the best energy and coordinates to find the minima, but also a small number of energy evaluations, as is shown in Figure 6.30.

Molecular Case for ARROBA: Water

This simple molecular system has been used as a starting point to test the algorithm. The results are shown in Table 6.21.

We started with the laziest geometry of 1.0000 Angstrom for the hydrogen-oxygen bond lengths and an HOH angle of 90°, which corresponds to the structure labeled as $\mathbf{q_2}$ (originally $\mathbf{q_1}$) in Table 6.21. As the algorithm starts the relabeling of the initial geometries is performed according to their relative energies. One of the main features

of ARROBA is shown when we note that both hydrogen-oxygen bond lengths are not identical, but that they become of the same length as the minimum is found.

Next, the goodness of the procedure is examined by comparing it with the steepest descent technique in Ab-initio and Semi-empirical program packages, in terms of the quality of the minimum found (geometry) and number of cycles required to converge to a minimal energey geometry. Table 6.22 shows results for the geometry optimization of water (all with the same input geometry) ACES II [83] (through a variety of Ab-initio basis sets), ZINDO and ARROBA. Also displayed is the experimental geometry.

Table 6.20. Variation on the number of energy evaluations (Cycles) with respect to the line search technique parameter α for the Approximate-ARROBA technique. Also displayed are the energy and coordinates of the minima found at each value of alpha. In the last, framed, line of the table we display the values of energy and coordinates that are expected for the minimum. All units are arbitrary.

α	Cycles	E(x,y)	X	Y
0.10	36	-0.5000	-0.7070	0.7069
0.15	22	-0.5000	-0.7071	0.7071
0.17	18	-0.5000	-0.7071	0.7071
0.18	16	-0.5000	-0.7071	0.7071
0.19	14	-0.5000	-0.7071	0.7071
0.20	12	-0.5000	-0.7071	0.7071
0.21	11	-0.5000	-0.7074	0.7071
0.22	15	-0.5000	-0.7072	0.7071
0.23	16	-0.5000	-0.7071	0.7071
0.25	14	-0.5000	-0.7069	0.7068
0.30	25	-0.4964	-0.7415	0.7086
Expected[a]		-0.5000	-0.7071	0.7071

a. This are the energy and coordinates accepted values for the minimum found. The other, symmetric, minimum for potential energy function is at (x,y) = (0.7071,-0.7071) with the same energy.

Table 6.22 indicates that despite the better overall geometry obtained by the Ab-initio calculations, particularly from the Double Zeta Polarization-Dunning basis set (DZP-DN), the semi-empirical geometry obtained by ZINDO is very good when compared with the experimental one. On the other hand, the optimized geometry coming from ARROBA is certainly encouraging.

We conclude that, despite the apparent extra cost of ARROBA because of the consecutive elimination of a given direction by perpendicular projection which implies more SCF calculations, its simplicity is appealing. The results from ARROBA are indeed encouraging, as it is well known that steepest descent works well when going downhill for steep slopes, but that along the valley of the minima shows a numerical disadvantage known as the zigzagging across the valley ground line [13, 88]. This seems no to be the problem with ARROBA because of its construction.

Table 6.21. Cycle-by-cycle downhill walk in the search for minimum of H_2O using the Approximate ARROBA technique. Energy (Hartrees), geometry (oxygen-hydrogen bond lengths (**r**, in Angstroms) and the hydrogen-oxygen-hydrogen angle (θ, in degrees)) and the maximum component of the gradient (g_{max}) at each cycle are displayed. The convergence criteria was that the displacement vector norm to be smaller than 10^{-3}.

	Energy	r_{OH2}	r_{OH3}	θ_{HOH}	g_{max}
q_1	-18.101384	1.0000	1.0800	90.00°	0.14890
q_2	-18.113190	1.0000	1.0000	90.00°	0.03385
q_3	-18.115546	0.9910	0.9902	92.84°	0.02926
q_4	-18.112815	1.0231	1.0279	92.25°	0.06088
q_5	-18.121658	0.9951	0.9956	102.77°	0.00806
q_6	-18.121647	0.9981	0.9966	102.71°	0.01186
q_7	-18.121975	0.9919	0.9929	104.71°	0.00917

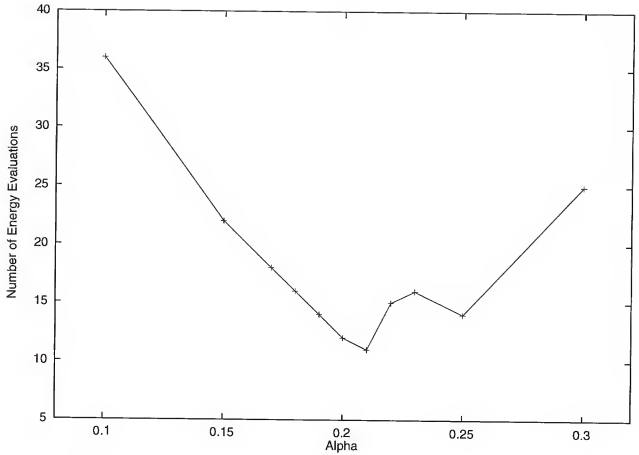


Figure 6.30. Number of energy evaluations required to find the minimum, as a function of the line search technique parameter α. The results corresponds to the approximate ARROBA technique applied to the potential energy surface with a minimum $E_{\mathbf{Min}}(x, y) = x^4 + y^4 + 2xy$.

Table 6.22. Equilibrium geometry for H_2O. We compare SCF results at different basis sets using the ACES II[a] (first six calculations) and ZINDO programs and compare them with experiment. Here the ARROBA results corresponds to the Approximate technique. **r** is the hydrogen-oxygen bond length (in Angstroms) and θ is the hydrogen-oxygen-hydrogen angle (in degrees)

	r	θ	Cycles
STO-3G	1.0135	97.28°	4
4-31G	0.9750	108.93°	4
6-31G*	0.9685	104.00°	4
6-31G**	0.9608	103.87°	4
DZP-DN	0.9624	104.44°	4
DZP-DIF	0.9639	105.02°	4
ZINDO	0.9952	106.74°	5
ARROBA	0.9924	104.71°	7
Experiment[b]	0.9573	104.50°	

a. ACES II is an Ab-initio program package whose main features were described in chapter 3.

b. Results taken from Tables 6.5 and 6.6 of reference [85].

CHAPTER 7
CONCLUSIONS AND FUTURE WORK

We have reviewed some of the most commonly used algorithms for Transition State (TS) and Geometry Optimization (GOPT) searches and their implementation in various available program packages (chapters 2 and 3, respectively), with discussion of their main features. We have pointed out their disadvantages, mainly related to cost (computer time) associated with the number of energy and/or Hessian evaluations required or to the fact that one needs to identify a suitable reaction coordinate. Other procedures need a large number of moves, or a good initial guess of the TS, whereas several others fail because of the large number of iterations required. In the case of geometry optimization, the problems, although fewer than for TS search, are similarly related to the size, initial guess, Hessian evaluations, and number of iterations required.

We suggest that for a model to be successful in the search for the TS, both R and P should be considered. The LTP procedures suggested here, including the Hammond Adapted ones, are based on this idea. For GOPT the LTP search features also are very convenient.

The main difference between the ideas suggested here and the procedures reviewed in this work, are based and defined by: the simultaneous consideration of products and reactants, a constrained step size, the simultaneous walking along a line connecting R and P towards the TS, the energy difference (ΔE) between the starting structures and a continuous downhill walk for GOPT, according to a zigzag type of search.

The concept of considering R and P approaching the TS simultaneously from both sides of the reaction is mainly based on the idea that each structure contains information and history about the other (having in common the same TS). The trajectory drawn in this way can be thought of as the lowest passage from one side of a mountain (R) to the other (P). The advantage of this idea is that we look to the mountain from both sides, and not from one. The GOPT algorithm was developed among the same lines, for which a second set of coordinates is automatically generated allowing for a fast and insured downhill walk.

The procedures investigated in this work are all characterized by the fact that no guess of a TS is needed and no evaluation of the second derivative matrix (Hessian) is required. This is the most important contribution of this work in comparison with the procedures reviewed in Chapter 2. Moreover, the Hammond-Adapted LTP (HALTP) procedures require a smaller number of calculations, still finding the TS with the same accuracy. Because of their simplicity they are very versatile. Avoiding the time consuming evaluation of the Hessian has been a general goal of this study.

Table 7.1. Summary of general properties and advantages of the Line-Then-Plane Technique for finding Transition States and to Optimize Geometry

Transition State search	*Geometry Optimization*
• No guess of **TS** is needed	• Only 1 initial geometry is needed
• R and P are needed	• Input can be a bad one
• Reduced number of calculations	• Simple down-hill walk
• Simple up-hill walk	• Reduced number of calculations
• Considers intermediate of reaction	• Updated Hessian procedure
• Path is easily approximated	
• Updated Hessian procedure	

Because of the advantages shown in Table 7.1, the procedures presented here will, at least in part, avoid some of the problems usually found in a TS search.

The procedures proposed in this work may fail only in cases characterized by a very steep reaction path. In those cases, it appears to be necessary to reconstruct more of the reaction path to ensure that the TS has been found and to reproduce the reaction path more accurately. This can be pursued by connecting the found first-order saddle point with reactants and products by means of a down-hill procedure, such as the steepest-descent technique (despite the inherent extra cost), especially if one is interested in kinetic aspects of the reaction. However, as has been advised at the beginning of this work, lower energy saddle points may exist (that can be found by means of several different procedures), as well as the many procedures to find them.

Although our procedures are simple and were constructed by inspection of a general potential energy surface, we have built up a strategy to find the TS that has the advantage of using a reduced number of calculations, has a simple and convenient writing of projected coordinates, updates the Hessian matrix, considers an intermediate of reaction, and involves the idea of finding the TS(s) starting simultaneously from R and P.

We believe that the new LTP TS and geometry optimization search procedures are well behaved and simple. It appears clear that more insight has to be obtained from the use of this procedures for the evaluation of bigger and more complicated molecular systems. We also foresee the inclusion of solvent effects in the TS searching, in order to analyze how the solvent influences the reaction mechanism.

On the other hand, work on an updating procedure of the projector is needed, mainly because as the initial guess in the perpendicular directions to the displacement vector is for now the projector itself. An *update* of it will speed up the Hessian update procedure in use.

For the geometry optimization procedure, we believe our results to be very promising. Some analysis is yet to be done on the line search parameter in order to construct a step for the perpendicular search adapted to the surface. On the other hand, it can be argued that ARROBA cannot compete with steepest descent techniques, however we remind that these techniques work very well for steep slopes but that they show a zigzagging behavior across an ample flat valley, which is a numerical disadvantage. We beleive that our only gradient technique, because of its projected features, overcomes this problem.

To accelerate convergence to maxima or minima, it can be argued that after a couple of movements, or when a change in the slope is detected, one should use the calculated projected points and fit them to a parabola to find the saddle points (the vertex of the parabola). This procedure has two inconveniences: first when we actually tried this in the model potential functions, the fitting gives the wrong answer either for maxima or minima mainly because the potential energy surface is somehow tortuous and not absolutely symmetric. Secondly, this fitting will eliminate the inherent beauty of the acceleration mechanisms of the LTP procedures.

Finally, it has become evident along the years, and according to the large amount of work devoted to tackle the search for maxima and minima, that chemical intuition in this kind of problems is always required. Unfortunately, if it fails one is lost in the hypersurface of algorithms with no saving recipe available. Consequently, one should handle it with care.

BIBLIOGRAPHY

[1] (a) H. W. Salzberg, *From Caveman to Chemist, Circumstances and Achievements*, American Chemical Society, Washington, DC 1991 ; (b) I. Asimov, *A Short History of Chemistry*, Anchor Doubleday & Co, New York, 1965 ; (c) B. Jaffe, *Crucibles, The Story of Chemistry*, Premier Book, New York, 1962.

[2] S. Arrhenius, Z. Physik. Chem. **4**, 226 (1889).

[3] H. Eyring and M. Polanyi, Z. Phys. Chem **B 12**, 279 (1931).

[4] Acc. Chem. Res. **28**, 119 (1995).

[5] C. Cardenas-Lailhacar and M. C. Zerner, to be published.

[6] G. S. Hammond, J. Amer. Chem. Soc. **77**, 334 (1955).

[7] A. H. Zewail and R. B. Bernstein, Chapter 9 in *The Chemical Bond, Structure and Dynamics* Ahmed Zewail, Ed. Academic Press Inc., San Diego California. 1992.

[8] A. Komornicki and J. W. McIver, Chem. Phys. Lett. **10**, 303 (1971).

[9] D. A. Liotard, Int. J. Quantum Chem. **44**, 723 (1992).

[10] H. B. Schlegel, Theor. Chim. Acta **83**, 15 (1992).

[11] P. Culot, G. Dive, V. H. Nguyen and J. M. Ghuysen, Theor. Chim. Acta **82**, 189 (1992).

[12] (a) J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **99**, 5257 (1993) ; (b) J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **99**, 5269 (1993) ; (c) J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **99**, 5276 (1993) ; (d) J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **101**, 2157 (1994) ; (e) K. Ruedenberg and J-Q. Sun, J. Chem. Phys. **101**, 2168 (1994)

[13] W. Quapp, Chem. Phys. Lett. **253**, 286 (1996) ; W. Quapp, Theor. Chim. Acta **75**, 447 (1989).

[14] C. Cardenas-Lailhacar and M. C. Zerner, Int. J. Quantum Chem. **55**, 429 (1995).

[15] (a) C. Gonzalez and H.B. Shlegel, J. Phys. Chem. **90**, 2154 (1989) ; (b) C. Gonzalez and H.B. Shlegel, J. Phys. Chem. **94**, 5523 (1990) ; (c) C. Gonzalez and H.B. Shlegel, J. Phys. Chem. **95**, 5853 (1991).

[16] J. W. McIver and A. Komornicki, J. Am. Chem. Soc. **94**, 2625 (1972).

[17] M. C. Zerner in Appendix C of A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, Mc Graw Hill, New York, USA, 1989.

[18] J. D. Head, B. Weiner and M. C. Zerner, Int. J. Quantum. Chem. **33**, 177 (1988).

[19] J. D. Head and M. C. Zerner in Adv. in Quantum Chem. **20**, 239 (1988). Also see reference [17] and M. Page, C. Doubleday and J. W. McIver Jr, J. Chem. Phys. **93**(8), 5634 (1990) For recent developments see: J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **99**, 5257 (1993) ; J-Q. Sun and K. Ruedenberg, J. Chem. Phys. **99**, 5269 (1993) ; J-Q. Sun, K. Ruedenberg and G. J. Atchity, J. Chem. Phys. **99**, 5276 (1993). For Symmetry breaking see: P. Sharfenberg, Theor. Chim. Acta. **53**, 279 (1979).

[20] B. A. Murtagh and R. W. H. Sargent, Comput. J. **13**, 185 (1970).

[21] C. G. Broyden, Math. Comput. **21**, 368 (1967); R. Fletcher, Comput. J. **13**, 23 (1970); Goldfarb, Math. Comput. **24**, 23 (1970); D. F. Shanno, Math. Comput. **24**, 647 (1970).

[22] R. Fletcher and M. J. D. Powell, Comput. J. **6**, 163 (1963) ; W.C. Davidon, Comput. J. **10**, 406 (1968); R. Fletcher, Comput.J. **8**, 33 (1965).

[23] J. Greenstadt, Math. Comput. **24**, 360 (1970).

[24] J.M. Bofill and M. Comajuan, J. Comp. Chem. **16**, 1326 (1995) ; J.M. Bofill, Chem. Phys. Lett. **260**, 559 (1996).

[25] (a) H. B. Schlegel, *Optimization of Equilibrium Geometries and Transition Structures*, in K. P. Lawley, editor of *Ab Initio Methods in Quantum Chemistry-I*, John Wiley & Sons Ltd., 1987, p. 249–286 ; (b) H. B. Schlegel, "Geometry Optimization on Potential Energy Surfaces", in Modern Electronic Structure Theory Part I, D. R. Yarkony Ed., Advanced Series in Physical Chemistry - Vol. 2., World Scientific, Singapore, Thailand, 1995. p: 459.

[26] M. C. Zerner, Department of Chemistry, University of Florida, Gainesville, Florida 32611. ZINDO, A General Semi-Empirical Program Package.

[27] J. D. Head and M. C. Zerner, Chem. Phys. Lett. **122**, 264 (1985).

[28] J. L. Kuester and J. H. Mize, "Optimization Techniques with Fortran", McGraw-Hill, Toronto, 1973.

[29] R. Fletcher, "Practical Methods of Optimization" (Wiley, Chichester, 1981).

[30] J.A. Nelder and R. Mead, Computer Journal **7**, 308 (1965). Also see: S.L.S. Jacoby, J.S. Kowalik and J.T. Pizzo *Iterative Methods for Nonlinear Optimization Problems*, Englewood Cliffs, N.J. Prentice Hall. 1972.

[31] K. Levenberg, Quant. Appl. Math. **2**, 164 (1994) ; D. W. Marquardt, SIAM J. 11, 431 (1963). NOTE: This was extended to minimization methods by Goldfeld et. al. and its behaviour in PES discussed by Simons et. al.: S. M. Goldfeld, R. E. Quandt and H. F. Trotter, Econometrica **34**, 541 (1966) ; J. Simons, P. Jorgensen, H. Taylor and J. Ozment, J. Phys. Chem. **87**, 2745 (1983).

[32] A. Banerjee, N. Adams, J. Simons and R. Shepard, J. Phys. Chem. **89**, 52 (1985).

[33] C. W. Gear, "Numerical Initial Value Problems in Ordinary Differential Equations", Prentice-Hall, Englewood Cliffs, NJ, 1971 ; K. Ishida, K. Morokuma and A. Komornicki, J. Chem. Phys. **66**(5), 2153 (1977) ; K. Muller and L. D. Brown, Theor. Chim. Acta **53**, 75 (1979) ; M. Page and J. W. McIver Jr., J. Chem. Phys. **88**, 922 (1988) ; M. Page, C. Doubleday and J. W. McIver Jr, J. Chem. Phys. **93**(8), 5634 (1990) ; D. Poppinger, Chem. Phys.Lett. **35**, 550 (1975).

[34] H. Hellmann, *"Einfuhrung in die Quantenchemie"* , F. Deuticke, Leipzig, 1937. R. P. Feynmann, Phys. Rev. **41**, 721 (1939). A. C. Hurley, Proc. Roy. Soc. **A226: 170**, 179 (1954).

[35] N. Metropolis, A. W. Rosenbluth, M. Rosenbluth, A. Teller, E. H. Teller, J. Chem. Phys. **21**, 1087 (1953) ; N. Metropolis et al., J. Chem. Phys. **21**, 1087 (1953) ; S. Kirkpatrick, C. D. Gelatt Jr. and M. P. Vecchi, Science **220**, 671 (1983) ; S. Kirkpatrick, J. Stat. Phys. **34**, 975 (1984) ; R. A. Donnelly, Chem. Phys. Lett. **136**, 274 (1987).

[36] M.P Allen and D.J. Tildesley *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1993).

[37] Th. A. Halgren and W. N. Lipscomb, Chem. Phys. Lett. **49 (2)**, 225 (1977).

[38] S. Ehrenson, J. Am. Chem. Soc. **96**, 3778 (1974); b) O.S. Tee, J. Am. Chem. Soc. **91**, 7144 (1969).

[39] A. Jensen, Theor. Chim. Acta **63**, 269 (1983).

[40] D. A. Liotard, E. F. Healy, J. M. Ruiz and M. J. S. Dewar, *"AMPAC Version 2.1"*, University of Texas at Austin, (1989) ; D. Liotard, *"Thesis"*, Universite de Pau, France (1979).

[41] M. J. S. Dewar, E.G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).

[42] M. J. S. Dewar, E. F. Healy and J. J. P. Stewart, J. Chem. Soc., Faraday Trans. II **80**, 227 (1984).

[43] Ch. Cerjan and W. H. Miller, J. Chem. Phys. **75 (6)**, 2800 (1981).

[44] H. B. Schlegel, J. Comput. Chem. **3**(2), 214 (1982).

[45] D. Poppinger, Chem. Phys.Lett. **35**, 550 (1975).

[46] D. J. Wales, J. Chem. Soc., Faraday Trans. **86**, 3505 (1990) ; D. J. Wales, J. Chem. Soc., Faraday Trans. **88**, 653 (1992); Also see: D. J. Wales, J. Chem. Soc., Faraday Trans. **86**, 3505 (1990) ; and D. J. Wales, J. Chem. Soc., Faraday Trans. **88**, 653 (1992)

[47] A. Komornicki and J. W. McIver, J. Am. Chem. Soc. **96**, 5798 (1974).

[48] B. H. Lengsfield III, J. Chem. Phys. **73**, 382 (1980).

[49] D. T. Nguyen and D. A. Case, J. Phys. Chem. **89**, 4020 (1985).

[50] D. R. Yarkony, Chem. Phys. Lett. **77**, 634 (1981) ; B. H. Lengsfield III and B. Liu, J. Chem. Phys. **75**, 478 (1981) ; R. Shepard, I. Shavitt and J. Simons, J. Chem. Phys. **76**, 543 (1982) ; P. Jorgensen, P. Swanstrom and D. L. Yeager, J. Chem. Phys. **78**, 347 (1983) ; J. T. Golab, D. L. Yeager and P. Jorgensen, Chem. Phys. **78**, 175 (1983) ; H. J. Werner, in "Ab Initio Methods in Quantum Chemistry", K. P. Lowley, Ed. Wiley, New York, 1987. p: 1 ; R. Shepard, in "Ab Initio Methods in Quantum Chemistry", K. P. Lowley, Ed. Wiley, New York, 1987. p: 63 ; H. J. Aa. Jensen, H. Agren and J. Olsen, in "Modern Techniques in Computational Chemistry: MOTECC-91", E. Clementi, Ed., ESCOM, Leiden, 1991. p: 599 ; J. D. Head and M.

C. Zerner in Adv. in Quantum Chem. **20**, 239 (1988) ; J. D. Head, B. Weiner and M. C. Zerner, Int. J. Quantum. Chem. **33**, 177 (1988) ; Y. G. Khait, A. I. Panin and A. S. Averyanov, Int. J. Quantum Chem. **54**, 329 (1995).

[51] H. J. Aa. Jensen and P. Jorgensen, J. Chem. Phys. **80**, 1204 (1984)

[52] D. K. Hoffman, R. S. Nord and K. Ruedenberg, Theor. Chim. Acta **69**, 265 (1986).

[53] P. Jorgensen, H. Jorgen, A. Jensen and T. Helgaker, Theor. Chim Acta **73**, 55 (1988).

[54] J. Pancir, Collect Czech. Chem. Commun **40**, 1112 (1975).

[55] K. Muller, Angew. Chem. Int. Ed. Engl. **19**, 1 (1980).

[56] M. V. Basilevsky, Chem. Phys. 67, 337 (1982) ; M. V. and A. G. Shamov, Chem. Phys. **60**, 347 (1981).

[57] N. Shida, J. E. Almlof and P. F. Barbara, Theor. Chim. Acta. **76**, 7 (1989).

[58] M. Dupuis, A. Farazdel, S. P. Karna and S. A. Maluendes in *HONDO: A General Atomic and Molecular Electronic Structure System* Chapter 6 of *MOTECC, Moddern Techniques in Computational Chemistry*, edited by E. Clementi (Escom, Leiden, The Netherlands, 1990), p. 277.

[59] K. Fukui, J. Phys. Chem. **74**, 4161 (1970).

[60] K. Fukui, Acc. Chem. Res **14**, 363 (1981) ; K. Ishida, K. Morokuma and A. Komornicki, J. Chem. Phys. **66**(5), 2153 (1977) ; A. KomornickiK. Ishida and K. Morokuma, Chem. Phys. Lett. **45**, 595 (1977) ; M. W. Schmidt, M. S. Gordon and M. Dupuis, J. Am. Chem. Soc. **107**, 2585 (1985)

[61] J. Andres, V. Moliner, J. Krechl and E. Silla, J. Phys. Chem. **98**, 3664 (1994).

[62] K. Muller and L. D. Brown, Theor. Chim. Acta **53**, 75 (1979).

[63] Y. Abashkin and N. Russo, J. Chem. Phys. **100**, 4477 (1994) ; Y. Abashkin, N. Russo and M. Toscano, Int. J. Quantum Chem. **52**, 695 (1994).

[64] K. Muller, Angew. Chem. **19**, 1 (1980).

[65] M. C. Zerner, *Semiempirical Molecular Orbital Methods*. In K. B. Lipkowitz and D. B. Boyd. editors of *Reviews in Computational Chemistry*, Vol. II, Chapter 8. VCH Publishing, New York, 1991, p. 313–366.

[66] E. Hückel, Z. Phys. **70**, 204 (1931).

[67] R. G. Parr *Quantum Theory of Molecular Electronic Structure*. Benjamin, New York, 1962; P. Lykos and R. G. Parr, J. Chem. Phys. **24**, 1166 (1956); R. Pariser and R. G. Parr, J. Chem. Phys. **21**, 466 (1953); J. A. Pople, Trans. Faraday Soc. **49**, 1375 (1953).

[68] M. J. S. Dewar, J. Am. Chem. Soc. **74**, 3341 (1952). A Molecular Orbirtal Theory of Organic Chemistry I. General Principles.; J. Am. Chem. Soc. 3345. A Molecular Orbital Theory of Organic Chemistry II. The Structure of Mesomeric Systems.; J. Am. Chem. Soc. 3350. A Molecular Orbital Theory of Organic Chemistry III. Charge Displacements and Electromeric Substituents.; J. Am. Chem. Soc. 3353. A Molecular Orbital Theory of Organic Chemistry IV. Free Radicals.; J. Am. Chem. Soc. 3353. A Molecular Orbital Theory of Organic Chemistry V. Theories of Reactivity and the Relationship Between Them.; J. Am. Chem. Soc. 3357. A Molecular Orbital Theory of Organic Chemistry VI. Aromatic Substitution and Addition.

[69] M. J. S. Dewar and A. J. Harget, Proc. Royal Soc. London Ser. A **315** 433 (1970).

[70] J. A. Pople, D. Santry and G. A. Segal, J. Chem. Phys. **43**, S129 (1965); J. A. Pople and G. A. Segal, J. Chem. Phys. **43**, S136 (1965).

[71] R. C. Bingham, M. J. S. Dewar and D. H. Lo, J. Am. Chem. Soc. **97**, 1285 (1975); R. C. Bingham, M. J. S. Dewar and D. H. Lo, J. Am. Chem. Soc. **97**, 1307 (1975).

[72] D. N. Nanda and K. Jug, Theor. Chim. Acta **57**, 95 (1980); D.N. Nanda and K. Jug, Theor. Chim. Acta **57**, 131 (1980); K. Jug and J. Schulz, J. Comput. Chem. **9**, 40 (1988); K. Jug and J. Schulz, J. Comput. Chem. **9**, 51 (1988).

[73] M. J. S. Dewar and W. J. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977).

[74] J. Baker, J. Comput. Chem. **7**, 385 (1986).

[75] W. Thiel, J. Am. Chem. Soc. **103**, 1413 (1981).

[76] M. Kolb and W. Thiel, Beyond the MNDO Model: Methodical Considerations and Numerical Results. In press.

[77] J. P. Stewart, *Quantum Chem. Prog. Exch.*, Prog. 445 (1983). MOPAC.

[78] J. J. P. Stewart, J. Comput.—Aided Mol. Design **4**, 1 (1990). MOPAC; J. J. P. Stewart, J. Comput. Chem. **10**, 209 (1989); J. J. P. Stewart, J. Comput. Chem. **10**, 221 (1989).

[79] J. D. Head and M. C. Zerner, Chem. Phys. Lett. **131**, 359 (1986).

[80] *Gaussian 88*. M. J. Firsch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. Martin, L. R. Kahn, J. J. P. Stewart, E. M. Fluder, S. Topiol and J. A. Pople. Gaussian Inc., Pittsburgh, PA, 1988.

[81] J. Baker, J. Comput. Chem. **8**(5), 563 (1987).

[82] C. Wentrup, "Reactive Molecules, The Neutral Reactive Intermediates in Organic Chemistry" (John Wiley & Sons, New York, 1984).

[83] J.F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale and R. J. Bartlett, Quantum Theory Project, University of Florida, 1993. ACES II, Version 1.0.

[84] J.A. Pople, D.L Beveridge and P.A. Dobosh, J. Chem. Phys. **47**, 2026 (1967) ; A.D. Bacon and M.C. Zerner, Theoret. Chim. Acta **53**, 21 (1979).

[85] A. Szabo and N. Ostlund, *Modern Quantum Chemistry*, Mc Graw Hill, New York, USA, 1989.

[86] Y. Abashkin and N. Russo, J. Chem. Phys. **100**(6), 4477 (1994)

[87] S. Bell and J. Crighton, J. Chem. Phys. **80**(6), 2464 (1984).

[88] W.H. Press, B.P Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York, USA, 1986.

BIOGRAPHICAL SKETCH

Cristian is the second son of Eduardo Cardenas-Diaz and Helene Lailhacar-Rochett. He was born in October the 25th, 1957, in Santiago, Chile.

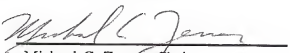
In his earliest memories he recalls the life in the countryside farm of his grandfather Luis Lailhacar-Lapeyre (who died too soon for Cristian, who continues to miss him), where he began with his sister and brother, to love nature, the wind, wide open spaces, life. At a very early age, he became interested in small creatures; insects soon became a passion. At 13 years old, he started to study entomology in the National Museum of Natural History in Santiago. Biology was then very appealing but not physics, and of course mathematics was out of the question. Although chemistry was not an early passion, he always felt a particular interest in atoms and their structure. Consequently, from the early studies of Leonardo Da Vinci, his hero by default, the history of chemistry, and the atomic bomb soon became an irresistible attraction. Scientists such as Albert Einstein, Max Planck, Madame Curie, Heisenberg, Ehrenfest, Desiderio Papp (a Chilean scientist), and others started filling his mind with fascinating ideas and their stories.

All this turned young Cristian to chemistry, but the implicit good background in physics and mathematics required was too much and, moreover, his high school grades in these subjects were not just bad, but terrible. With almost a colossal effort he overcame this natural barrier, studied chemistry and soon became an assistant professor of calculus. He decided to write his master's thesis in quantum chemistry

at the University of Chile. This opened unimaginable doors for Cristian, who married Alejandra, his girlfriend of more than 10 years. Then in 1991 he came to the University of Florida to work on his Ph.D. in chemistry, under the scientific supervision of Dr. Michael C. Zerner, a good friend.

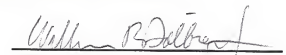
Life has been good to Cristian, who has found his marriage blessed with two beautiful daughters, Francisca and Catalina, and his loving wife Alejandra. Friends and family have always been a vital part of his life and they have always been there for him and his family. Being a Catholic, he has always felt a deep devotion to the church, but for national (Chilean) historical reasons, it has always been the Virgin of El Carmen, patroness of Chile, to whom he asked for help in hard times and his prayers for everything in life were devoted.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



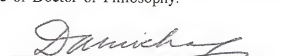
Michael C. Zerner, Chair
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



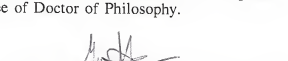
William Dolbier
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



David A. Micha
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



N. Yngve Öhrn
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Samuel B. Trickey
Professor of Physics

This dissertation was submitted to the Graduate Faculty of the Department of Chemistry in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1998

Dean, Graduate School